Titre: Title:	Diagnosis of Machining Conditions Based on Logical Analysis of Data
Auteur: Author:	Yasser Shaban
Date:	2014
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Shaban, Y. (2014). Diagnosis of Machining Conditions Based on Logical Analysis of Data [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/1578/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/1578/
Directeurs de recherche: Advisors:	Soumaya Yacout, & Marek Balazinski
Programme: Program:	Génie industriel

UNIVERSITÉ DE MONTRÉAL

DIAGNOSIS OF MACHINING CONDITIONS BASED ON LOGICAL ANALYSIS OF DATA

YASSER SHABAN

DÉPARTEMENT DE MATHÉMATIQUES ET DE GENIE INDUSTRIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION

DU DIPLÔME DE PHILOSOPHIAE DOCTOR

(GÉNIE INDUSTRIEL)

NOVEMBRE 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

DIAGNOSIS OF MACHINING CONDITIONS BASED ON LOGICAL ANALYSIS OF DATA

présentée par : SHABAN Yasser

en vue de l'obtention du diplôme de : Philosophiae Doctor

a été dûment accepté par le jury d'examen constitué de :

M. ADJENGUE Luc-Désiré, Ph.D., président

Mme YACOUT Soumaya, D.Sc., membre et directeur de recherche

M. BALAZINSKI Marek, Docteur ès Sciences, membre et codirecteur de recherche

M. ZDZISLAW Klim, Ph.D., membre

M. TIAN Zhigang, Ph.D., membre

DEDICATION

"Whoever follows a path in the pursuit of knowledge, GOD will make a path to paradise easy for him."

-Muhammad (pbuh)

"Simplicity is the ultimate sophistication."

-Leonardo da Vinci

"As complexity rises, precise statements lose meaning and meaningful statements lose precision."

-Lofti Zadeh

"The more knowledge you accumulate, the more critical thinking you gain."

-Soumaya Yacout

ACKNOWLEDGMENTS

I am greatly indebted to my supervisor Soumaya Yacout. I feel grateful to her for taking time out of her busy schedule to answer my many questions. Without her guidance and her dedication, I doubt I would have been able to complete my Ph.D and write this thesis in such a proficient and timely manner. I admit that she is the best teacher ever I have had in my life and I feel honored to be one of her students. She simply knows how to bring out the best of me. She truly inspired me to work with sincerity and dedication, and even changed the trajectories of my thoughts.

I would like to express my deep thanks to my co-supervisor Marek Balazinski, for his support, encouragement, and helpful suggestions. He has been helping and guiding me to the right direction to complete my thesis.

I am also thankful to Luc-Désiré Adjengue, Zhigang Tian, Klim Zdzislaw and Jolanta Sapieha for agreeing to attend my Ph.D. defense as jury members and representative of the Dean.

I would like to express my deep appreciation to my colleagues for their help and supporting. I consider them not only colleagues, but valuable friends.

I am heavily indebted to my family and specially my wife for their understanding, encouragement, and perseverance in spite of the circumstances that obliged us and being away during my Ph.D.

RÉSUMÉ

Un élément clé pour un système d'usinage automatisé sans surveillance est le développement de systèmes de surveillance et de contrôle fiables et robustes. Plusieurs modèles mathématiques et statistiques, qui modélisent la relation entre les variables indépendantes et les variables dépendantes d'usinage, sont suggérés dans la littérature, en commençant par le modèle de Taylor jusqu'aux modèles de régression les plus sophistiqués. Tous ces modèles ne sont pas dynamiques, dans le sens que leurs paramètres ne changent pas avec le temps. Des modèles basés sur l'intelligence artificielle ont résolu de nombreux problèmes dans ce domaine, mais la recherche continue. Dans la présente thèse, je propose l'application d'une approche appelée Analyse Logique de Données (LAD) pour prédire le sortant d'un processus d'usinage. Cette approche a démontré une bonne performance et des capacités additionnelles une fois comparée à la conception traditionnelle des expériences ou à la modélisation mathématique et statistique. Elle est aussi comparée dans cette thèse à la méthode bien connue des réseaux de neurones. Elle est basée sur l'exploitation des données saisies par des capteurs et l'extraction des informations utiles à partir de ces dernières. LAD est utilisé pour déterminer les meilleures conditions d'usinage, pour détecter l'usure de l'outil, pour identifier le moment optimal de remplacement de l'outil d'usinage, et pour surveiller et contrôler les processus d'usinage. Étant donné que les capteurs et les technologies de l'information sont tous les deux en expansion rapide et continue, il serait prévu qu'un outil d'analyse tels que LAD aidera à tracer un chemin dans l'amelioration des processus d'usinage en utilisant les techniques de pointe afin de réduire considérablement le coût ces processus. Les résultats de mon travail pourraient avoir un impact important sur l'optimisation de ces processus.

ABSTRACT

A key issue for an unattended and automated machining system is the development of reliable and robust monitoring and controlling systems. Research in Artificial Intelligence-based monitoring of machining systems covers several issues and has solved many problems, but the search continues for a robust technique that does not depend on a statistical learning background and that does not have ambiguous procedures. In this thesis, I propose the application of an approach called Logical Analysis of Data (LAD) which is based on the exploitation of data captured by sensors, and the extraction of useful information from this data. LAD is used for determining the best machining conditions, detecting the tool wear, identifying the optimal replacement time for machining tools, monitoring, and controlling machining processes. LAD has demonstrated good performance and additional capabilities when it is compared to the famous statistical technique, Proportional Hazard Model (PHM), and the well known machine learning technique, Artificial Neural Network (ANN).

Since sensors' and information technologies are both expanding rapidly and continuously, it is expected that an analysis tool such as LAD will help in blazing a new trail in machining processes by using state of the art techniques in order to significantly reduce the cost of machining process.

TABLE OF CONTENTS

DEDICATION……………………………………………… ……………………………iii

ACKNOWLEDGMENTS iv

RÉSUMÉ v

ABSTRACT vi

TABLE OF CONTENTS vii

LIST OF TABLES xi

LIST OF FIGURES xiii

LIST OF ABBREVIATIONS xvi

INTRODUCTION 1

General objectives 1

Logical Analysis of Data 3

Originality of research 3

Literature review 4

CHAPTER 1 THESIS ORGANIZATION 9

CHAPTER 2 DIAGNOSIS OF MACHINING OUTCOMES BASED ON MACHINE LEARNING WITH LOGICAL ANALYSIS OF DATA 12

2.1 Summary 12

2.2 Introduction 12

2.3 Introduction to Logical Analysis of Data 14

2.4 Methodology: Logical Analysis of Data 15

2.5 Experiment and Results 18

2.6 Learning and validation 21

2.7 Discussion 23

CHAPTER 3 ARTICLE 1: PROCESS CONTROL BASED ON PATTERN RECOGNITION FOR ROUTING CARBON FIBER REINFORCED POLYMER 28

3.1 Abstract 29

3.2 Introduction 29

3.3 Experiment Description 31

3.4 Logical analysis of data (LAD) 35

3.4.1 The methodology 35

3.4.2 Numerical Example 38

3.5 Peformance comparison 41

3.5.1 The ANN technique 41

3.5.2 Validation and comparison 43

3.6 Process control system 45

3.7 Conclusion 52

CHAPTER 4 ARTICLE 2: OPTIMAL REPLACEMENT OF TOOL DURING TURNING TITANIUM METAL MATRIX COMPOSITES 53

4.1 Abstract 54

4.2 Introduction 54

4.3 Model description 55

4.4 Optimal replacement Policy 57

4.5 Decision rule 58

4.6 Experiment description 58

4.7 Development the model and results 62

4.8 Conclusion 68

Identifying optimal replacement times for cutting tools 69

CHAPTER 5 ARTICLE 3: OPTIMAL REPLACEMENT TIMES FOR MACHINING TOOL DURING TURNING TIMMCS UNDER VARIABLE MACHINING CONDITIONS.............. 80

5.1 Abstract 81

5.2 Introduction 81

5.3 Model description of a tool operating in varying conditions 83

5.4 Optimal replacement Policy 85

5.5 Description of the Experiment 86

5.6 Development the model and results 90

5.7 Practical use and sensitivity analysis 95

5.8 Conclusion..97

CHAPTER 6 ARTICLE 4: TOOL WEAR MONITORING AND ALARM SYSTEM BASED ON PATTERN RECOGNITION WITH LOGICAL ANALYSIS OF DATA............................. 98

6.1 Abstract 99

6.2 Introduction 99

6.3 Logical analysis of data (LAD)...101

6.4 Description of the Experiment 104

6.5 Knowledge extraction and learning...106

6.6 The Statistical Proportional Hazards Model (PHM)...109

6.7 LAD on-line alarm system development and comparison with the PHM alarm function...113

6.8 Discussion and Conclusion 117

CHAPTER 7 ARTICLE 5: TOOL REPLACEMENT BASED ON PATTERN RECOGNITION WITH LOGICAL ANALYSIS OF DATA.. 121

7.1 Summary and conclusion 122

7.2 Introduction 122

7.3 Experiment description 123

7.4 Logical Analysis of Data (LAD)..127

7.5 Optimal replacement & decision rule..128

7.5.1 Pattern-based Optimization. 129

7.5.2 Statistical-based Optimization...130

7.6 Decision Rule 132

CHAPTER 8 ARTICLE 6: CUTTING TOOL WEAR DETECTION USING MULTI-CLASS LOGICAL ANALYSIS OF DATA 136

8.1 Abstract 137

8.2 Introduction 137

8.3 Logical analysis of data (LAD)...139

8.4 Experiment description and wear classification...143

8.5 Results and discussion..149

8.5.1 Validation and comparison...151

8.6 Conclusion..154

GENERAL DISCUSSION.. 156

CONCLUSION 158

REFERENCES... 159

LIST OF TABLES

Table 2.1: The experimental results 20-21

Table 2.2: Pattern found when, force and torque are considered 25

Table 3.1: A sample of the experimental results 33-34

Table 3.2: collected Boolean observation vectors and their classes 39

Table 3.3: The remaining dataset after founding the first positive pattern 40

Table 3.4: ANN architectures for the four models. 43

Table 3.5: Accuracy of the ANN and the LAD techniques. 45

Table 3.6: The positive patterns obtained by LAD for problems (A-1) and (A-2), and the negative patterns obtained for problems (B-1) and (B-2). 46

Table 3.7: Two runs of the simulated Process control using LabVIEW 50-51

Table 4.1: The experimental results showing the wear of tool number 6 61

Table 4.2: Times to failure and samples of wear value inspections for the six tools 62

Table 4.3: Summary of goodness of fit test results 63

Table 4.4: Summary of cost analysis 65

Table 4.5: Summary of availabilty analysis 66

Table 4.6: The experimental results showing the wear of tool number 6 74

Table 4.7: Summary of cost analysis 76

Table 4.8: Summary of availabilty analysis 77

Table 5.1 : The coded design of experiment 87

Table 5.2: The design of experiment 87

Table 5.3:The experimental results showing the wear of tool 1-1 89

Table 5.4: Times to failure *TTF* for the 28 tools 90

Table 5.5: Summary of goodness of fit test results 92

Table 6.1: Design of experiment 105

Table 6.2: Experimental data of tool 1, replication2 107-108

Table 6.3: Worn patterns 108

Table 6.4: Times to failure for the 28 tools 110

Table 6.5: Summary of goodness of fit test results 112

Table 6.6: Replacement decision for PHM-model and LAD alarm 116

Table 7.1:Experiment details....124

Table 7.2: Experimental data of tool number 9 126

Table 7.3: Times to failure for the 10 tools 127

Table 7.4: replacement time for 10 tools based on LAD 130

Table 7.5: replacement times comparison for ten tools 135

Table 8.1: Experimental data of cut 5 (tool 1, $ap = 1.5$ mm) 148

Table 8.2: patterns found by the software cbmLAD 149

Table 8.3: machining variables' frequencies and class identifiers in each of the five classes 150

Table 8.4: the results of testing procedure 152

LIST OF FIGURES

Figure 1-1: Thesis organization....10

Figure 2-1: Experimental setup showing: (a) the fixture font plate, (b) the fixture back plate, (c) the force dynamometer and (d) the high speed slip ring, (Rawat and Attia 2009a)....18

Figure 2-2 : The observed damages by delamination at the end of tool life: (a) entry delamination, v = 15,000 rpm, (b) entry delamination, v = 12,000 rpm,....19

Figure 2-3: Observation in (a) entry and exit delamination, (b) hole circularity at entry, (c) hole circularity at exit, (d) hole diameter error at entry, (e) hole diameter error at exit, and (g) hole surface roughness....24

Figure 3-1: Trends for the feed force (Fx), transverse force (Fy) , and axial force (Fz) for different speeds, feeds and tool overhang length (TL1 = 38mm) (Meshreki et al. 2012)....32

Figure 3-2: Schematic of the experimental setup....32

Figure 3-3: ANN models: (A) controllable variables model, (B) monitored uncontrollable variables model....42

Figure 3-4: The machining process control....47

Figure 3-5: On-line machining process control for delamination quality using LabVIEW....49

Figure 3-6: Flow chart of process control....49

Figure 4-1: The experiment setup....59

Figure 4-2: Tool wear measurements for 6 tools....59

Figure 4-3: Wear interpolating....60

Figure 4-4: Condition-based replacement policy-cost analysis....64

Figure 4-5: Condition-based replacement policy-availability analysis....66

Figure 4-6: Condition-based replacement policy-optimal decision....67

Figure 4-7: Tool wear measurements for 6 tools....73

Figure 4-8: Condition-based replacement strategy-optimal decision....78

Figure 5-1:The experimental setup....87

Figure 5-2:Wear interpolating 88

Figure 5-3:Summary of estimated parameters (based on ML method). 91

Figure 5-4: logarithmic reliability function plot for each run 92

Figure 5-5: Hazard rate curves for each run 93

Figure 5-6: Optimal replacement function-cost analysis 93

Figure 5-7: Optimal replacement function-availability analysis 94

Figure 5-8: Finding the optimal replacement time in cost and availability analysis 95

Figure 5-9:Optimal replacement example-cost analysis 96

Figure 5-10: The cost ratio sensitivity 97

Figure 6-1: Forces directions during turning 105

Figure 6-2: Schematic diagram of experimental setup 106

Figure 6-3: Wear classification and failure threshold 107

Figure 6-4: Summary of estimated parameters (based on ML method) 111

Figure 6-5: The model with only the significant variables (the best model to be used) 112

Figure 6-6: Schematic diagram for LAD on-line alarm system 114

Figure 6-7: On-line alarm system front panel. 115

Figure 7-1: The experiment setup 124

Figure 7-2: Tool wear measurements for 10 tools 125

Figure 7-3: Tool wear identification on statistical-based and pattern-based 129

Figure 7-4: Replacement decisions 133

Figure 8-1: Six cuts with different cutting variables 144

Figure 8-2: recursion steps for DP algorithm 145

Figure 8-3: The feed forces and the cutting forces vs. the tool wear at six combinations of feed rate and depth of cut representing the six cuts. 146

Figure 8-4: Wear classes classification 147

Figure 8-5: ANN architecture 153

LIST OF ABBREVIATIONS

AE	Acoustic Emission
ANN	Artifical Neural Networks
CFRP	Carbon Fiber Reinforced Polymer
CMM	coordinate measuring machine
CNC	Computer Numerical Control
EFE	Equivalent Fixture Error
KSTest	Kolmogorov-Smirnov test
LAD	Logical Analysis of Data
MILP	Mixed Integer Linear Programming
ML	Maximum Likelhood
MMC	Metal Matrix Composites
PAHD	Physical Asset Health Diagnosis
PHM	Proportional Hazard Model
SMOT	Synthetic Minority Over-sampling Technique
SVM	Support Vector Machine
TTF	Time to Failure
TTT	Total Time on Test plot

INTRODUCTION

Currently, manufacturing enterprises must compete in a global market with growing demands for better quality, greater choices of products with shorter products' life-cycles, and reduced costs. Machining process exists in almost every manufacturing company. For many products, the machining process constitutes the critical step that determines whether products conform to the quality specifications defined by the product design. Controlling this process is thus fundamental. We use knowledge discovery technique based on pattern recognition in order to monitor and control a machining process. We consider machining of composites material such as titanium metal matrix composites (TiMMCs) and Carbon Fiber Reinforced Polymer (CFRP). MMCs have light weight and high strength which are suitable for aerospace industries in order to improve the performance of an aircraft. Despite being expensive, MMCs have become viable in various fields such as biomedical and aerospace industries. CFRP is an important composite material. It has many applications in aerospace and automotive fields. We consider the three principal machining processes namely, turning, drilling and milling.

General objectives

In this doctoral research, the first objective is the implementation of Logical Analysis of Data (LAD) on a machining process by monitoring the outputs of some experimental trials. By evaluating specific qualities such as surface roughness, delamination, hole circularity error in drilling, the machining conditions that lead to conforming products are found. By determining which conditions lead to conforming products or non-conforming products, a machining process mapping based on LAD is built. The measurements of machining forces are used to evaluate predict the quality and geometric profile of the machined part. Therefore, force monitoring is used in the diagnosis of the part accuracy. The measurements of forces are used to evaluate the quality and geometric profile of the machined part. Two-class LAD model is applied in order to determine the best conditions for that machining process.

We then propose a new process control technique, and we apply it to a routing process. The measured machining conditions are used to evaluate the quality and the geometric profile of the machined part. The machining conditions, whether controllable (time independent) or uncontrolled (time dependent) are used to control part accuracy and its quality by using LAD. By detecting conjunctions of threshold values and characteristic patterns for these conditions, we use them to control the quality of a machined part at a specific accepted range. LAD is used in order to find the characteristic patterns that lead to conforming products and those which lead to nonconforming products. It is used for online control of a simulated routing process of CFRP. We apply it to the high speed routing of woven carbon fiber reinforced epoxy, and we compare the accuracy of LAD to that of an Artificial Neural Network (ANN), since it is probably the most known machine learning technique. By using experimental results, and based on the simulated model, we show how LAD is used to control the routing process by tuning autonomously the routing conditions.

The second objective of this thesis is finding the optimal tool replacement when the cutting tool's condition is degrading. First, we find the tool replacement time when a tool is used under constant machining conditions, namely the cutting speed, the feed rate, and the depth of cut, during turning MMCs. The Proportional Hazard Model (PHM) is used to model the tool's reliability and hazard functions using Exakt software. Experimental data are obtained and used to construct and validate the PHM model, which is then used in decision making. Second, we find the optimal time to tool replacement when the tool is used under variable machining conditions, namely the cutting speed, and the feed rate. PHM is used to find an optimal replacement function. Third, a new tool wear monitoring and alarm system that is based on LAD is presented. The system is a non-intrusive on-line device that measures the cutting forces and relates them to tool wear through learned patterns. It is developed during turning MMCs. Fourth, we show how to exploit condition monitoring data in machining operation in order to extract intelligent knowledge, and use this knowledge to determine the tool replacement time. We compare the results obtained when applying LAD to those obtained by using the well-known statistical PHM considering multi-objective optimization.

In an automated manufacturing environment, monitoring of tool wear in the machining process is essential for avoiding tool failure, increasing machine utilization, and decreasing production cost. The third objective is finding the status of tool wear by applying a multi-class model to a machining process. By using experimental data, tool wear classes are found using Douglas-Peucker algorithm. LAD is then used as a knowledge discovery technique to find a hidden correlation between machining variables which leads to detect and identify wear classes.

Logical Analysis of Data

LAD is a data mining technique that can classify phenomena based on pattern recognition. LAD is applied in two consecutive phases: a learning or training phase, and a testing or theory formation phase, in which part of the database is used to extract special features or patterns of some phenomena, and the rest of the database is used to test the accuracy of the previously extracted knowledge. LAD is based on supervised learning; this means that the database contains its classes. LAD was first proposed in (Crama et al. 1988). After many years, LAD become one of the most promising data mining methods developed to date for extracting knowledge from data (Han et al. 2011).

In 2007, LAD was used in the field of industrial engineering for the first time (Salamanca and Yacout 2007). After that, many studies researched the use of LAD in engineering applications. LAD methodology has certain advantages, such as: interpretability power and causality identification. These make LAD very useful in addressing engineering problems. It has no restriction on the type of data. This makes LAD capable of handling different types of data simultaneously such as event data, condition monitoring data, or both. LAD is a non-statistical approach, thus there is no need to make certain assumptions regarding the posteriori class probabilities.

Originality of research

The originality and novelty of research is as follows:

1. To the best of our knowledge, LAD has never been used in machining process applications for fault Diagnosis.
2. A new method for machining process control based on LAD's patterns is presented. The developed process controller can be used to control the machining process by tuning autonomously the machining conditions.
3. A new method for tool wear monitoring alarm system is developed. The results show that tool wear monitoring with alarm system gives an accurate alarm for cutting tool replacement.

4. The tool wear monitoring with alarm system is developed during turning titanium metal matrix composites (TiMMCs). These composites are a new generation of materials which have proven to be viable in various industrial fields such as biomedical and aerospace, and they are very expensive.

5. We showed how to exploit condition monitoring data in machining operation in order to extract intelligent knowledge, and use this knowledge to determine the cutting tool replacement time.

6. A new tool wear multi-class detection method is presented. Based on experimental data, tool wear classes are defined using Douglas-Peucker algorithm, thus LAD is used in a situation of unsupervised learning, which is, to our knowledge, a first.

7. Process Controller and Alarm System for Machining Operations (Pro-CASMO) is developed. Pro-CASMO is an invention for tool wear monitoring and machining process controller. It is composed of two modules: on-line tool wear monitoring with alarm system, and on-line machining process controller. cbmLAD (c. Software 2012), a platform of PXI, and LabVIEW software were used to develop Pro-CASMO. By using the experimental results obtained under sequential different machining conditions, the on-line tool wear monitoring with alarm system (Pro-CASMO_Module#1) is developed. By using the experimental results, and based on a simulated model, the on-line machining process controller (Pro-CASMO_Module#2) is used to control the machining process by tuning autonomously the machining conditions.

Literature review

Previous research in the diagnosis of machining conditions has covered several topics and many possible solutions, but the search continues for a robust technique that does not depend on statistical learning and does not have ambiguous procedures. For example, in (Ho-Cheng and Dharan 1990) , the authors used a fracture mechanics approach to find the optimal thrust force that reduces delamination and improves the product quilty while drilling composite materials. This result does not take into consideration the possible interaction between the different force and the torque. In (Chen 1997), the author concluded that thrust force and torque are different with and without the onset of delamination for the drilling of CFRP composite laminates. In addition, they found a step linear relationship between the delamination and the average thrust force for drilling unidirectional CFRP composite laminates with a carbide drill. (Rawat and Attia 2009b, 2009a) introduced the concept of machinability maps of woven carbon fiber composites under high speed drilling conditions (up to 15,000 rpm) and established the effect of cutting conditions on the quality characteristics of drilled holes, namely delamination, geometric errors, and surface finish. They concluded that thrust and cutting force have considerable effect on quality maps. This work was further extended in (Attia 2011.) in order to cover a speed range of up to 40,000 rpm. In (Zuperl et al. 2012), the authors used neural networks and fuzzy logic to control the cutting force in the process of ball-end milling, and in order to maintain constant roughness. In (Bustillo and Correa 2012), the authors studied the optimization of roughness in deep drilling operations under high speed conditions. The cutting force, the cutting parameters, and the cooling system were considered input variables to Bayesian Networks. In (Tansel et al. 2013), the authors used Torque-based Machining Monitor to estimate the remaining tool life and to detect the chatter from the torque signal. They concluded that this procedure was a good choice for monitoring procedure particularly when multiple spindles work simultaneously on the same work piece.

Many Researches have been conducted on milling process control. They used artificial intelligence learning techniques in order to control machining process. For example, in (P. B. Huang 2014), The author developed an intelligent neural-fuzzy model for surface roughness monitoring system in milling operations. He developed a decision-making system which analyzed the cutting forces and then responded with an accurate output. He concluded that his developed system can be used, in future, as an adaptive control system of the machining parameters in smart Computer Numerical Control (CNC) machine. In (Zhang et al. 2007), the authors used ANN to develop surface roughness adaptive control in turning process. They used data from controllable cutting parameters such as feed rate, cutting speed, and depth of cut, and also uncontrollable monitored parameters such as vibration signals in order to develop neural-networks-based surface roughness adaptive control system. Other researchers used other techniques, for example, (Coker and Shin 1996) used ultrasonic sensing to control surface roughness during machining processes. (H. Wang and Huang 2006) used the concept of an Equivalent Fixture Error (EFE) to improve machining process control. Based on simulated data, they illustrated their concept. In (Du et al. 2012), the authors developed a robust approach for root causes identification in machining process using hybrid learning algorithm and engineering-driven rules. In order to judge whether the process is in normal or abnormal condition, off-line pattern match relationships and on-line time series measurements were used. They validated the developed approach by using data from the real-world cylinder head of engine machining processes.

Much research tried to improve tool life in several ways. For example, Klim et al (Klim et al. 1996) proposed a method to improve cutting tool life in machining using the effect of feed variation on tool wear and tool life. By changing feed rate, the reliability function is changed, and thus the tool life is changed. The Weibull distribution was used to fit the data. The experiment was conducted under constant cutting speed. Balazinski and Mpako (M Balazinski and Mpako 2000) proposed an improvement of tool life through using two discrete feed rates. The method depends on varying the feed rate throughout the cutting process. By varying the feed, the tool-chip contact area increases, the tool wear rate decreases and consequently leads to improvement of the cutting tool life. The experiment was conducted under constant cutting speed. Lin and Shyu (Lin and Shyu 2000) concluded that using variable feed machining, and constant cutting speed, when drilling stainless steel is a significant method for improving the cutting tool life.

Other researches tried to find the optimal replacement strategy by using PHM for modelling tool life, then using another technique to find optimal strategy. For example, Mazzuchi and Soyer (Mazzuchi and Soyer 1989) used a PHM to assess machine tool reliability. Fully Bayesian analysis is used to find optimal machining conditions. Liu and Makis (H. Liu and Makis 1996) derived a formula to calculate the cutting tool reliability under variable cutting conditions. They used PHM while considering the machining conditions as covariates. In(P. H. Liu et al. 2001), the work was extended by developed algorithm based on stochastic dynamic programming for finding the optimal tool replacement times in a flexible manufacturing system. Ding and He (Ding and He 2011) used a PHM by considering vibration signals as a time–dependent covariate. The author suggests that vibration signals are good indicators to tool wear. Reliability analysis based on feature extraction from tool vibration signals is introduced. They found remarkable relationship between the tool condition monitoring information and the life distribution of tool wear by using PHM. Other research used classical Weibull distribution to fit tool life distribution. For example, In (Vagnorius et al. 2010), the Weibull distribution is used to fit tool life distribution. The optimal replacement time for metal cutting is determined from a total time on test (TTT) plot. Some researchers tried to improve the cutting tool life by changing feed rates while the cutting speed is constant (Klim et al. 1996; M Balazinski and Mpako 2000; Lin and Shyu 2000), others consider the PHM as good model for tool life representation(Mazzuchi and Soyer 1989; V Makis 1995; Tail et al. 2010). In most of these models, it was assumed that the machining conditions have significant effect over the entire tool life but finding tool replacement models is still unavailable.

The tool wear in machining processes is analyzed by two approaches: Firstly, theoretical and numerical approach, such as state space methods and finite element method (FEM), and secondly, data-driven approach, such as artificial neural network (ANN) and fuzzy logic (Shi and Gindy 2007). Li (Li 2012) presented an exclusive review of tool wear estimation using theoretical analysis and numerical simulation technologies. Sick (Sick 2002) presented an exclusive review of indirect online tool wear monitoring in turning with ANN as an example of data-driven technique. By indirect, we mean that researchers usually measure covariates (variables) which are indirectly correlated with tool wear such as the cutting forces. These forces are measured on-line during machining process. There are hundreds of researches about tool wear monitoring system. Nevertheless, only a few systems found their way to real industrial application (Jemielniak 1999). The tool wear monitoring systems development is still on-going attempt (Sick 2002). Byrne et al (Byrne et al. 1995) presented a review about utilization of these systems in industry. Another review about commercial tool monitoring systems was done by (Jemielniak 1999).

Due to the availability of sensory signals, data-driven approach has received much attention to build on-line tool wear monitoring systems. Data-driven techniques need training stage to learn how to adjust adaptively to the data without statistical distribution. Once learning stage is accomplished and validated, the system can detect worn pattern correctly. (Damodarasamy and Raman 1993) developed an inexpensive system for classifying tool wear states using pattern recognition. Despite that the accuracy of classification was relatively small, they concluded that pattern recognition can be successfully used to predict the status of cutting tool wear. They combined the feed force, radial force and the root mean square of acoustic emission (AE) signals to predict the tool wear. In (Shi and Gindy 2007), the tool wear predictive model is presented by combination of least squares support vector machines and principal component analysis technique. The platform of PXI and LabVIEW were used to develop the system. (S Purushothaman and Srinivasa 1994) developed a model for classifying a worn-out tool and a fresh tool. They used ANN for building a model. (Kang et al. 2007) developed a method of pattern recognition of tool wear based on discrete hidden Markov models. The results showed that the proposed method is effective. All techniques which used pattern recognition for classifying tool wear states are based on assumptions related to the data structure. In this work, the proposed technique, LAD is not based on any assumptions or statistical techniques. It is used for the first time in tool condition monitoring. In this paper, our objective is to report and discuss the results obtained experimentally.

The tool wear detection in machining processes is estimated by two approaches: Continuous tool wear estimation and tool wear classification (Sick 2002). Researchers estimate tool wear continuously using data driven techniques e.g. (Marek Balazinski et al. 2002; Achiche et al. 2002; Ren et al. 2008). (Sick 2002) presented an exclusive review of online tool wear detection in turning and found that the majority of researches considered tool wear detection based on tool wear classification. For example, (Damodarasamy and Raman 1993) used three adjacent classes of tool wear in order to develop a detection tool wear model using pattern recognition. They concluded that pattern recognition can be used to detect the classes of cutting tool wear. (S Purushothaman and Srinivasa 1994) developed tool wear monitoring model by using two classes, worn-out tool and a fresh tool. They used Artificial Neural Networks (ANN) for building the model. (Ertunc and Oysu 2004) used five classes to develop tool wear monitoring system using dynamic hidden Markov models. (Kang et al. 2007) developed tool wear monitoring model using pattern recognition based on discrete hidden Markov models. A three classes cutting tool wear model is used. (Tobon-Mejia et al. 2012) used five classes in order to diagnose the wear's progression. They used dynamic Bayesian networks' technique. Although the continuous tool wear estimation gives a better picture of the tool wear progression, many researchers consider that in practical situations the tool wear classification is quite sufficient for allowing the operator to make an informed decision (Sick 2002).

CHAPTER 1 THESIS ORGANIZATION

The research is divided in three parts as shown in figure 1-1. The research began by implementing LAD methodology on machining process, in part one. The two topics that are introduced are : diagnosis of machining process, and controlling of machining process. In chapter 2, LAD is used to characterize the effect of cutting forces on the quality of a machined part made of CFRP material. LAD is used in order to map the machining conditions, in terms of force and torque that lead to conforming products and those which lead to nonconforming products. In chapter 3, a new process control technique is introduced in routing process of CFRP. The measured machining conditions are used to evaluate the quality and the geometric profile of the machined part. The machining conditions are then used to control part accuracy and its quality. LAD is used to finds the characteristic patterns which lead to conforming products and those that lead to nonconforming products. These patterns are used to control the quality of a machined part at specific range. By using experimental results, and based on a simulated model, we showed how LAD is used to control the routing process by tuning autonomously the routing conditions.

As the research progressed, the problem of finding the optimal replacement time and replacement decision making is introduced in part two. This part consists of four chapters. Finding the tool replacement time when a tool is used under constant machining conditions, namely the cutting speed, the feed rate, and the depth of cut, during turning TiMMCs is introduced in chapter 4. Proportional Hazard Model (PHM) is used to model the tool's reliability and hazard functions using Exakt software. Experimental data are obtained and used to construct and validate the PHM model, which is then used in decision making. In chapter 5, the problem evolved, we found the optimal time to tool replacement when the tool is used under variable machining conditions, namely the cutting speed, and the feed rate. Two optimality models for cost minimization and availability maximization are introduced. In chapter 6, a new tool wear monitoring and alarm system that is based on LAD is introduced. The system is a non-intrusive on-line device that measures the cutting forces and relates them to tool wear through learned patterns. It is developed during turning TiMMCs. We showed that the tool life can be increased by giving an alarm at the right moment. The proposed monitoring system is tested by using the experimental results obtained under sequential different machining conditions and validated by comparing it to the limit obtained when the statistical Proportional Hazard Model (PHM) is used. We finished part two by showing how to exploit condition monitoring data in machining operations in order to extract intelligent knowledge, and use this knowledge to determine the cutting tool replacement time. The optimal tool replacement times were found by considering cost-availability optimization.

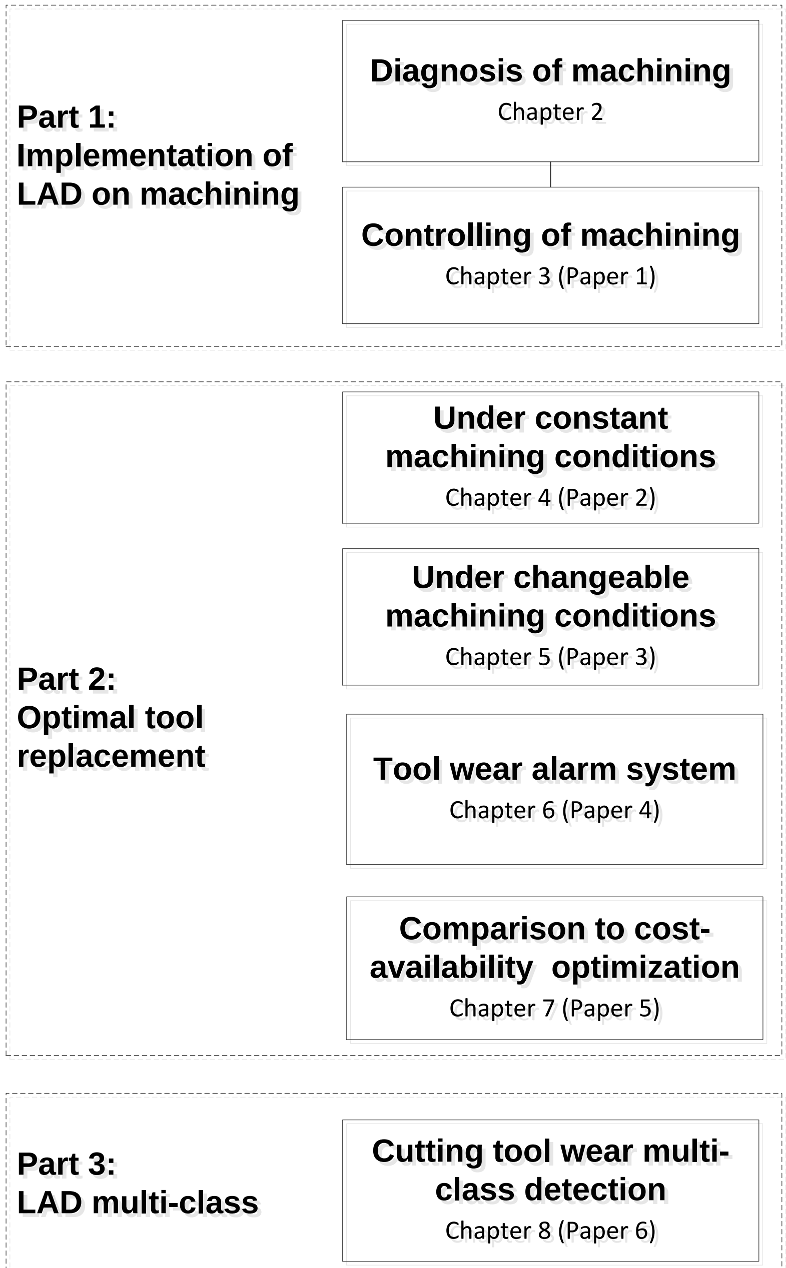


Figure 1-1: Thesis organization

In the last part, part 3, a new tool wear multi-class detection method is presented. By using experimental data, tool wear classes are found using Douglas-Peucker algorithm. Logical

Analysis of Data (LAD) is then used as a knowledge discovery technique based on unsupervised data, in order to find a hidden correlation between machining variables which leads to the detection and and the identification of wear classes.

CHAPTER 2 DIAGNOSIS OF MACHINING OUTCOMES BASED ON MACHINE LEARNING WITH LOGICAL ANALYSIS OF DATA

2.1 Summary

Force is considered to be one of the indicators that best describe the machining process. Measured force can be used to evaluate the quality and geometric profile of the machined part. In this chapter, a combinatorial optimization approach is used to characterize the effect of force on the quality of a machined part made of Carbon Fiber Reinforced Polymers (CFRP) material. The approach is called Logical Analysis of Data (LAD) and is based on machine learning and pattern recognition. LAD is used in order to map the machining conditions, in terms of force and torque that lead to conforming products and those which lead to nonconforming products. In this chapter, the LAD technique is applied to the drilling of CFRP plates, and the results, based on data obtained experimentally, are reported. A discussion of the potential use of LAD in manufacturing concludes the chapter.

2.2 Introduction

In some industries such as aerospace, automotive and aircraft structure, the composite materials are the backbone of manufacturing due to their structures or/and machinability characteristics (Rahman et al. 1999). The composite materials have unique mechanical properties, namely, high strength-to-weight ratio, high fracture toughness, and excellent corrosion resistance properties. Machining of composites, particularly drilling, is extensively used in the production of riveted and bolted joints. Any defect arising during machining has a significant technical and economic impact. For many products, drilling constitutes the critical process that differentiates between conforming and non-conforming products according to quality characteristics defined by the product design. Controlling this process is thus fundamental (Haber et al. 2002; Benardos and Vosniakos 2002; Liang et al. 2004). Drilling composite materials is more difficult than drilling metals in general because of the non-homogeneous composition and abrasive behaviour of reinforcing fibres. The tool confronts fibres and matrix , whose response to the machining process could be completely versatile (Teti 2002). As such, the process of drilling composites needs to be characterized, and the process' parameters optimized, in order to meet the design tolerances and to achieve defect-free components.

In (Ho-Cheng and Dharan 1990) , the authors used a fracture mechanics approach to find the optimal thrust force that reduces delamination while drilling carbon fiber-epoxy laminates. This result does not take into consideration the possible interaction between the different force and the torque. In (Chen 1997), the author concluded that thrust force and torque are different with and without the onset of delamination for the drilling of Carbon Fiber Reinforced Polymer (CFRP) composite laminates. In addition, they found a step linear relationship between the delamination and the average thrust force for drilling unidirectional CFRP composite laminates with a carbide drill. (Rawat and Attia 2009b, 2009a) introduced the concept of machinability maps of woven carbon fiber composites under high speed drilling conditions (up to 15,000 rpm) and established the effect of cutting conditions on the quality characteristics of drilled holes, namely delamination, geometric errors, and surface finish. They concluded that thrust and cutting force have considerable effect on quality maps. This work was further extended in (Attia 2011.) in order to cover a speed range of up to 40,000 rpm. In (Zuperl et al. 2012), the authors used neural networks and fuzzy logic to control the cutting force in the process of ball-end milling, and in order to maintain constant roughness. In (Bustillo and Correa 2012), the authors studied the optimization of roughness in deep drilling operations under high speed conditions. The cutting force, the cutting parameters, and the cooling system were considered input variables to Bayesian Networks. In (Tansel et al. 2013), the authors used Torque-based Machining Monitor to estimate the remaining tool life and to detect the chatter from the torque signal. They concluded that this procedure was a good choice for monitoring procedure particularly when multiple spindles work simultaneously on the same work piece. Many researchers consider force and torque as the indicators that best describe the machining process. Cutting force and torque are used in modelling, optimization, and controlling the machining process. In this chapter, LAD is used to study the effects of uncontrollable machining conditions, force and torque, on the quality characteristics of composite materials when they are subjected to a machining process. The advantage of LAD is its capacity to capture possible correlations between force and torque, and the quality of output. LAD is based on pattern recognition and classification. It is not based on statistical technique, which means that it does not impose any assumptions related to the nature or structure of the data. In the following section, we introduce LAD and we show how it is used in order to analyze and characterize the drilling process of carbon fiber reinforced material.

2.3 Introduction to Logical Analysis of Data

An important development in Physical Asset Health Diagnosis (PAHD) is the introduction of techniques for data analysis, diagnosis and prognosis that depend mainly on the continuous development in the field of information technology. As sensor technology advances, users are collecting more data than ever before. Special techniques are needed in order to extract useful information out of this data. With the advancement in computer technology, it is now possible to manipulate large volumes of data and to extract valuable knowledge out of it. LAD is an artificial intelligence technique that allows the classification of phenomena based on pattern recognition (Bennane and Yacout 2012). Characterization and classification are the major functions of artificial intelligence techniques (Choudhary et al. 2009) . LAD is applied in two consecutive phases, the learning or training phase, and the testing or the theory formation phase, where part of the database is used to extract special features or patterns of some phenomena, and the rest of the database is used to test the accuracy of the previously learned knowledge. We note that LAD is a technique based on supervised learning; this means that the database contains uncontrollable variables of the machining process and the corresponding output. In the process of drilling of composite materials, the output is the product's final quality measured with respect to predefined specifications, specifically, inner and outer delamination, inner and outer hole diameter error, inner and outer hole circularity, and surface roughness. In this chapter, the controllable variables, namely speed (v) and feed (f) are constant. Their effects have been already studied in (Rawat and Attia 2009b, 2009a) and in (Yacout et al. 2012). The monitored uncontrollable variables are thrust force (Fz), cutting force (Fc), and torque (M). After the accomplishment of the two phases of training and testing, characteristic patterns which represent condition thresholds on the machining variables are found by the LAD. These patterns are used in order to produce a map that divides the machining variables' space into conforming and nonconforming products. This map can later be used to monitor the uncontrollable variables, and according to their combined values, action can be taken in order to avoid the spaces that lead to nonconforming products and to stay in the spaces leading to conforming ones. The practical consequence of this finding is that the mapping technique can be used in condition monitoring in order to predict the quality of outcomes and to give an alarm if the machining process is going into the nonconforming spaces.

The idea of LAD was first proposed in (Peter L Hammer 1986; Crama et al. 1988). Research efforts have helped transform the approach into a methodology for data analysis and applications in the medical, industrial and economics fields (Alexe et al. 2007). To extract features from a database and to recognize patterns, LAD divides the database into two sets; the first set is used to extract special features or patterns of some phenomena, while the second part is used for the testing phase of the previously learned theory (Soumaya Yacout 2010). If the accuracy is not satisfactory, then some of LAD's parameters are changed, for example, the size of the training and testing data sets, the number of classes of outputs, the number and the nature of features, and the characteristics of the generated patterns. In (P.L. Hammer and Bonates 2006), a LAD overview is introduced by a group of researchers of RUTCOR at Rutgers University in USA. In (P.L. Hammer and Bonates 2006; Soumaya Yacout 2010) LAD methodology was compared to the most popular techniques of machine learning, such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), and other AI techniques. The same work had also been done in (Mortada et al. 2011). It was concluded that LAD is comparable, and in some cases outperformed comparable techniques.

2.4 Methodology: Logical Analysis of Data

LAD is used to extract knowledge from a dataset consisting of any type of observation; binary, numerical or nominal. Originally, LAD was used as a two-class (conforming π^+, nonconforming π^-) classification technique (Bores et al. 2000). The observations were classified as either positive (conforming) π^+(class 1) or negative (non-conforming) π^- (class 2), depending on whether they were observed when the process was producing non-conforming or conforming products. A special characteristic of LAD is the extraction of the collections of patterns which characterizes each class. These patterns represent interactions between variables (force and torque) in each class, positive or negative, separately. As such, the patterns are also called positive or negative depending on whether they describe phenomena found in the positive π^+ or in the negative π^- set of observations. LAD is used as a pattern-based classifier if new observations that are not included in the original dataset need to be classified (Bores et al. 2000). In (E. Mayoraz and Moreira 1997) and (Moreira 2000) there are two approaches on the extension of LAD to multi-class applications. In those works, the different methods that break down a multi-class classification problem (Polychotomy) into two-class problems (Dichotomies) are described. LAD is used as a multi-class diagnostic technique for the detection and identification of faults (M.-A. Mortada et al. 2013). In this chapter we describe and apply the two- class LAD only.

The main steps of the LAD are the binarization of data, pattern generation, and theory formation. Data binarization is the process of transformation of a database of any type into a Boolean database. There are many techniques for data binarization and research in this field is abundant (Eddy Mayoraz and Moreira 1999). The binarization of a continuous numerical feature A, and the number of resulting binary attributes that replace it, are dependent on the number of distinct values of A in the training data set. The binarization procedure used in this chapter starts by ranking, in ascending order, all the distinct values of the numerical feature A as follows:

$$u_A^{(1)} < u_A^{(2)} < \cdots < u_A^{(Q)} \qquad (Q \leq M) \qquad (1)$$

Where Q is the total number of distinct values of the feature A, and M is the total number of observations in the training set. The cut-points $\varepsilon_{A,j}$, where j is the number of cut points for each feature, are found between each pair of values that belong to different classes. By averaging these two values as shown in equation (2) the cut-points are calculated as follows:

$$\varepsilon_{A,j} = (u_A^{(k)} + u_A^{(k+1)})/2 \qquad (2)$$

Where $u_A^{(k)} \in \pi^+$ and $u_A^{(k+1)} \in \pi^-$ or vice versa. A binary attribute b is then formed from each cut-point. Each cut-point $\varepsilon_{A,j}$ has a corresponding binary attribute $b_{\varepsilon_{A,j}}$ with defined value:

$$b_{\varepsilon_{A,j}} = \begin{cases} 1 & \text{if } u_A \geq \varepsilon_{A,j} \\ 0 & \text{if } u_A < \varepsilon_{A,j} \end{cases} \qquad (3)$$

The patterns generation procedure is the key building block in the LAD decision model. A positive pattern is defined as a conjunction of some binary attributes, which is true for at least one positive observation, and false for all negative observations in the training data set. A negative pattern is defined similarly. The number of binary attributes used to define the pattern is called the degree of a pattern d. For example, pattern p of degree d is a conjunction of d attributes. That pattern covers an observation in the training set if and only if it is true for that particular observation (Bores et al. 2000). As such, if a pattern covers an observation, this means that this observation is from the same class to which the pattern belongs. There are many techniques for pattern generation, for example enumeration (Bores et al. 2000), heuristics (Peter L Hammer 1986; P.L. Hammer and Bonates 2006), and linear programming (Ryoo and Jang 2009).

Theory formation is the final step in the LAD decision model. A discriminant function is formulated in equation (4) in order to generate a score ranging between -1 and 1. When the output of a discriminant function is a positive value that means that the tested observation O belongs to the positive class, and negative otherwise. Zero value means no classification is possible (M.-A. Mortada et al. 2011).

$$\Delta(O) = \sum_{i=1}^{N^+} \alpha_i^+ Z_i^+(O) - \sum_{i=1}^{N^-} \alpha_i^- Z_i^-(O) \qquad (4)$$

where $N^+(N^-)$ is the number of positive (negative) patterns, $Z_i^+(O)\left(Z_i^-(O)\right)$ is equal to 1 if pattern (i) covers observation O, and is equal to zero otherwise, $\alpha_i^+(\alpha_i^-)$ is the weight of the positive (negative) pattern $p_i^+(p_i^-)$.

For each new observation O, the calculated value $\Delta(O)$ varies between +1 and -1, where +1(-1) is an indication of the domination of the positive (negative) patterns, hence an indication that the observation belongs to the conforming (non-conforming) space. A zero or near zero value means that the observation cannot be classified in either class.

Two measures of accuracy, ACCURACY and ν (the quality of classification) are used.

$$\nu = \frac{a+d}{2} + \frac{e+f}{4} \qquad (5)$$

Where the values (a) and (d) represent the proportion of observations, positive and negative, which are correctly classified. The values (e) and (f) represent the proportion of observations, positive and negative, which are not classified. Another measure is the ACCURACY $= \frac{C+B}{Mt}$, where C is the total number of correctly classified positive observations, B is the total number of correctly classified negative observations, and Mt is the total number of observations in the testing set.

2.5 Experiment and Results

In this section LAD is used to characterize the machining process of Carbon Fiber Reinforced Polymers (CFRP) material in terms of force and torque. The experimental results presented in Table I were reported in (Rawat and Attia 2009b, 2009a) for a quasi-isotropic laminate comprising of 28 plies of woven graphite epoxy. The laminate was manufactured by autoclave molding with a cure time of 60 min at 260 F under 516.75 kPa autoclave pressure to produce a final cured thickness of 6.35 mm. A two-flute, 5 mm diameter drill with a 30o helix angle, 118o point angle, a fluted length of 44.5 mm, and a total length of the drill of 76.2 mm was used in the tests. The carbide grade of the drill used was ISO K10-K20 with approximately 7% cobalt as binder. The laminate was sandwiched between the front and back plates of the machining fixture shown in Figure 2-1.

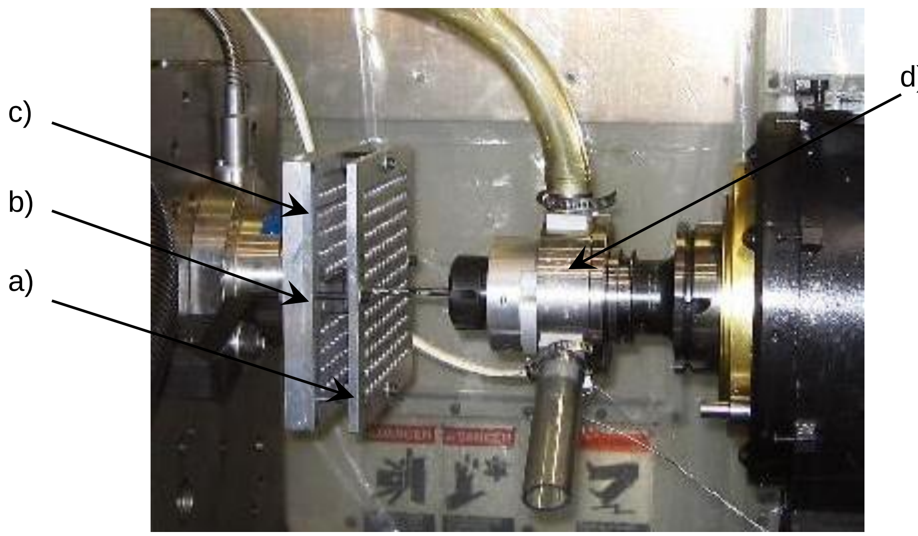


Figure 2-1: Experimental setup showing: (a) the fixture font plate, (b) the fixture back plate, (c) the force dynamometer and (d) the high speed slip ring, (Rawat and Attia 2009a)

The experiment is performed at five different speeds (v) with seven different feeds (f), which results in 35 observations as summarized in Table 1. In each observation, the thrust force (Fz), the cutting force (Fc) and the torque (M) are monitored and measured. The force was recorded using a 9272-type Kistler dynamometer as shown in Figure 1. At each one of 35 observations, delamination at the hole entry and exit, hole circularity at entry and exit, hole diameter error at entry and exit, and surface roughness are measured. These characteristics of the machined part represent the quality of the drilling process (S. Yacout et al. 2012; Rawat and Attia 2009a).

Delamination at the hole entry and exit are defined by the parameter η in terms of the hole diameter D_h and the maximum delamination damage diameter Dm (Fig. 2-2(b)), where $\eta = (D_m - D_h)/D_h$. Figure 2-2 shows the delamination at entry for different speeds.

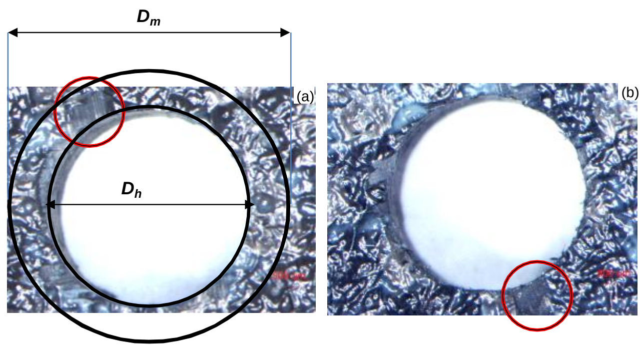


Figure 2-2 : The observed damages by delamination at the end of tool life: (a) entry delamination, v = 15,000 rpm, (b) entry delamination, v = 12,000 rpm,

The delamination is analyzed using the Olympus Model GZX 12 optical microscope. Hole circularity and diameter error are measured using a ''Mitutoyo-Mach 806'' coordinate measuring machine (CMM). Surface topography measurements were done in accordance with the International Standard IS0 4288:1996 using a Form Talysurf series 2 surface profilometer. The specifications of these qualities are as follows:

Exit and entry delamination = 1.0

Hole circularity at exit and at entry ≤ 0.2%

Hole diameter error at exit and at entry ≤ 0.02%

Hole surface roughness ≤ 0.5 µm.

The set of observations that satisfies (doesn't satisfy) any of these specifications is called positive (conforming, Π^{+}) (negative (non-conforming, Π^{-}))

Table 2.1: The experimental results

Variables or attributes						Quality outcomes						
No	v rpm 10^3	f micron/ rev	Fz N	Fc N	M N.mm	Entry Delam.	Exit Delam.	Circul. in % at Entry	Circul. in% at Exit	Dim Error in% at Entry	Dim Error in% at Exit	Surface rough. µm
1	1.5	20	58.2	26.4	131.9	1	1	0.056	0.074	0.016	0.032	0.45
2	5	20	54.3	20.8	103.8	1	1	0.128	0.112	0.02	0.138	1.19
3	8.5	20	47	17.3	86.7	1	1	0.142	0.118	0.044	0.182	2.25
4	12	20	36.5	14.7	73.6	1	1	0.402	0.202	0.072	0.088	3.28
5	15	20	30.4	13.1	65.5	1.08	1.07	0.484	0.594	5.65	4.8	2.03
6	1.5	60	77.5	39.4	196.8	1	1	0.072	0.108	-0.104	-0.006	0.6
7	5	60	59	34.7	173.6	1	1	0.154	0.122	-0.104	-0.004	0.9
8	8.5	60	46.5	28.4	142.1	1	1	0.158	0.136	-0.082	-0.026	0.72
9	12	60	43.5	22	109.9	1	1	0.17	0.122	-0.052	0.038	0.82
10	15	60	48	25.6	127.8	1	1	0.104	0.094	0.156	0.064	1.06
11	1.5	100	78.5	53.1	265.5	1	1	0.086	0.168	-0.114	-0.084	0.97
12	5	100	60	45	224.8	1	1	0.132	0.156	-0.104	-0.05	1.16
13	8.5	100	67	43.5	217.7	1	1	0.156	0.152	-0.098	-0.034	0.99
14	12	100	62.3	38.2	191	1	1	0.172	0.136	-0.068	0.014	1.2
15	15	100	50	31.9	159.5	1	1	0.176	0.116	0.21	0.004	1.14
16	1.5	200	103	64.1	320.3	1.25	1.08	0.128	0.136	-0.138	-0.052	1.43
17	5	200	98	63.1	315.3	1.22	1.05	0.134	0.166	-0.12	-0.06	1.31
18	8.5	200	93	50.8	254.2	1.22	1.05	0.164	0.2	-0.064	-0.04	1.82
19	12	200	90	53.1	265.5	1.21	1.04	0.178	0.144	-0.064	-0.046	1.89

Table 2.2: The experimental results (continued)

20	15	200	103	58.6	292.9	1.16	1.03	0.154	0.13	0.1	0.034	1.87
21	1.5	400	200	108.2	540.8	1.29	1.11	0.16	0.216	-0.164	-0.074	1.62
22	5	400	175.7	91	454.8	1.24	1.06	0.166	0.158	-0.11	-0.05	1.76
23	8.5	400	162	87.2	435.8	1.25	1.06	0.182	0.22	-0.068	-0.07	2.24
24	12	400	154	73	365.2	1.23	1.06	0.198	0.158	-0.048	-0.09	2.14
25	15	400	140.2	71.1	355.7	1.16	1.05	0.208	0.136	0.09	-0.036	2.27
26	1.5	600	370	141.3	706.5	1.33	1.22	0.172	0.25	-0.284	-0.116	2
27	5	600	310	111.6	558.2	1.3	1.14	0.188	0.208	-0.144	-0.09	1.78
28	8.5	600	260	99.5	497.3	1.27	1.1	0.204	0.152	-0.124	-0.136	2.78
29	12	600	182.5	75.9	379.3	1.26	1.04	0.238	0.162	-0.09	-0.094	2.26
30	15	600	145.5	87.9	439.6	1.16	1.05	0.312	0.152	0.03	-0.024	1.91
31	1.5	800	570	164.8	823.9	1.48	1.24	0.178	0.198	-1.004	-1.09	2.49
32	5	800	461.4	140.8	704.1	1.4	1.17	0.413	0.453	-0.968	-1.09	1.96
33	8.5	800	310	118.4	591.8	1.38	1.11	0.556	0.244	-0.15	-0.068	3.04
34	12	800	184	96.6	483	1.37	1.07	0.648	0.202	-0.084	-0.042	2.94
35	15	800	147	76	380.1	1.35	1.05	0.756	0.184	0.036	-0.022	2.2

2.6 Learning and validation

Our objective is to use the data presented in Table 1 to train LAD to detect automatically and without human interference, the threshold values and characteristic patterns for machining conditions that lead to acceptable quality of drilled parts, and those that lead to unacceptable quality. In order to reach this objective the software cbmLAD (c. Software 2012) is trained by using the data obtained from the experimental results shown in Table 3.1. To study the effects of the uncontrollable variables on the quality of outcomes, and find the corresponding characteristic patterns, only force and torque, i.e. the monitored variables are considered. The effects of speed and feed on the quality of outcomes were presented in (S. Yacout et al. 2012). These allow for the prediction of the quality of the outcome by monitoring force and torque only over time. For each one of the seven quality characteristics, specifications divide the outcomes' space into two distinct spaces, one for conforming products and one for nonconforming products. Set O of the thirty five observations is also divided into two sets of training, *L*, and testing, *T*. In this chapter, we present the results obtained when the training set is composed of 34 observations and testing is formed of the remaining observation. The process is repeated 35 times where each observation was chosen exactly once to constitute the testing set. The quality of classification is calculated as given in section 2. The results show that ν = 99.4 %. The only wrong classification happened when the data point was unique, that is the knowledge contained in the observation that constitutes the test set was not repeated in any other observation. For example, from table 1, it can be seen that the specification for surface roughness is satisfied in only one observation out of thirty five, at the combination (spindle speed=1500 rpm, feed=20 micron/rev). Obviously, when this observation constituted the testing set, the classification result was wrong since the training set, and consequently the learning process, did not contain any acceptable quality of surface roughness (Kohavi 1995). This training and testing procedure is known as leave-one-out (or jackknifing) cross validation procedure, which is considered by many machine learning references as the best validation procedure when the amount of data for training and testing is limited (Ian H Witten and Frank 2011). This procedure is attractive for two reasons: first, the greatest possible amount of data is used for training, which presumably increases the chance that the classifier is an accurate one. Second, the procedure is deterministic: no random sampling is involved.

We also conducted another well-known cross validation procedure which is the tenfold. In this case, the data is divided randomly into 10 parts in which each class is represented in approximately the same proportion as in the full dataset. Each part is held out in turn and the learning process is trained on the remaining nine-tenths; then its error rate is calculated on the holdout (or testing) set. Thus, the learning procedure is executed a total of 10 times with different training sets. The results show that ν = 94.6 %.

2.7 Discussion

Table 2.2. exhibits the patterns found by the software cbmLAD (c. Software 2012). The indicators are the uncontrollable variables: thrust force, cutting force and torque. The patterns that lead to conforming products are called positive, and those that lead to non-conforming products are called negative. The target is to map the conditions that lead to conforming products and those that lead to non-conforming products, in terms of force and torque. Patterns are generated for all seven qualities as shown in table 2. For example, the generated positive pattern for delamination illustrates the threshold boundary for conforming (positive) parts as (33.45<Thrust force, Fz<84.25) which means that as long as thrust force is between these thresholds, the machined part will be conforming to the required specification for delamination. This unique pattern is found in all positive observations from number 1 to number 4 and from number 6 to 15. We note that the negative observations 5, and 16 to 35, are not covered by this pattern. The patterns generated for nonconforming (negative) parts are twofold: the negative pattern number 1 (Thrust force,Fz >84.25), which covers the observations from number 16 to number 35, and the negative pattern number 2, (Thrust force,Fz < 33.45) which covers the observation number 5. In these two identified regions, machined parts are expected to be nonconforming to the delamination specifications. For the delamination quality characteristic, it may seem easy for anyone to distinguish between the positive and negative observation by looking at table 1, and by using only one attribute (Thrust force,Fz), but the generated map problem is more fragmented when we consider the others qualities since the regions where the products are conforming or nonconforming are not obviously separable. Nevertheless, cbmLAD identifies and characterizes these regions perfectly and by using the lowest possible number of attributes. We also emphasize that the more uncontrollable variables we have, and the more our database increases in size, the more it will be impossible to do the mapping manually or graphically, thus the importance of using LAD.

As an illustration, we plot the positive (●) and negative (○) observations, which are obviously non separable as shown in Figure 3-3, for entry and exit delamination 2-3(a), hole circularity at entry 2-3(b) , hole circularity at exit 2-3(c), hole diameter error at entry 2-3(d), hole diameter error at exit 2-3(e) and hole surface roughness 3-3(g).

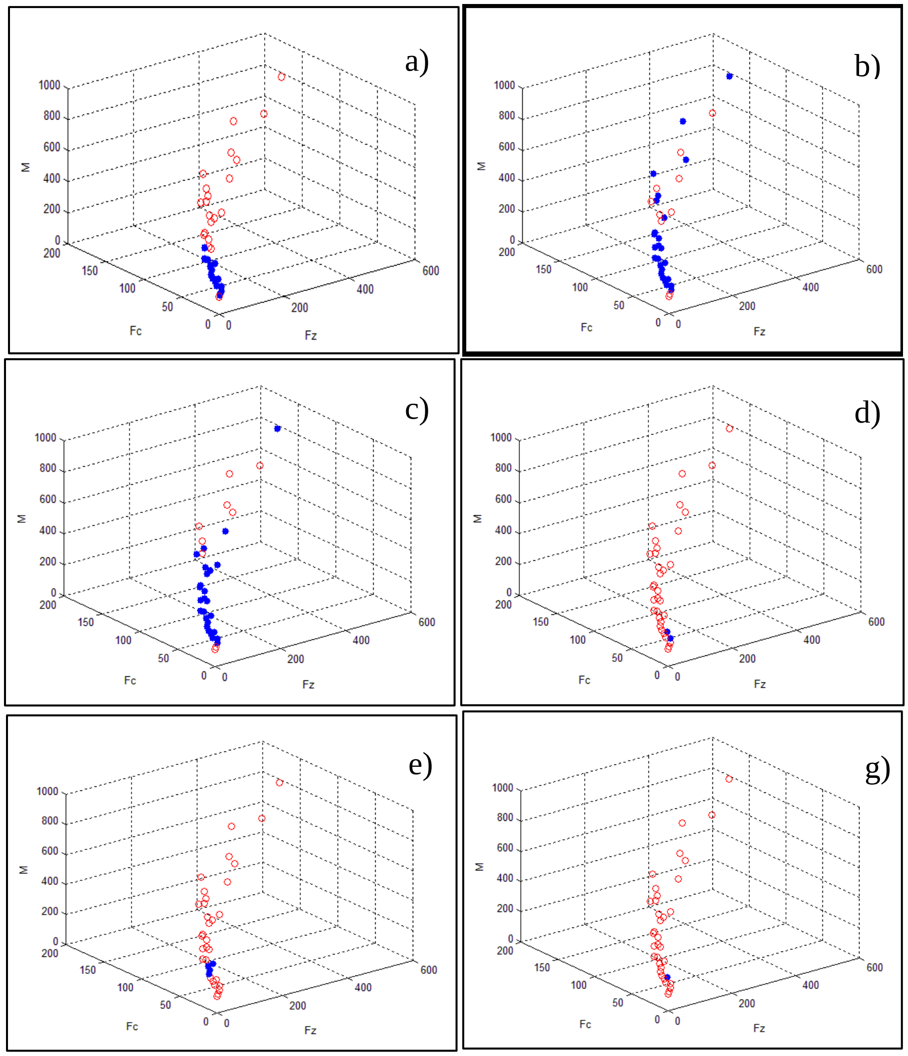


Figure 2-3: Observation in (a) entry and exit delamination, (b) hole circularity at entry, (c) hole circularity at exit, (d) hole diameter error at entry, (e) hole diameter error at exit, and (g) hole surface roughness

Table 2.3: Pattern found when, force and torque are considered

Positive patterns for delamination (exit and entry)	
Pattern 1:	
Thrust force	Greater Than 33.45
Thrust force	Less Than 84.25
Negative patterns for delamination (exit and entry)	
Pattern 1:	
Thrust force	Greater Than 84.25
Pattern 2:	
Thrust force	less Than 33.45
Positive patterns for hole circularity at exit	
Pattern 1:	
Thrust force	Greater Than 40
Torque	Less Than 407.95
Pattern 2:	
Cutting force	Greater Than 87.55
Torque	Less Than 468.9
Pattern 3:	
Thrust force	Greater Than 515.7
Pattern 4:	
Thrust force	Greater Than 230
Thrust force	Less Than 285
Negative patterns for hole circularity at exit	
Pattern 1:	
Thrust force	Less Than 515.7
Cutting force	Greater Than 103.805
Pattern 2:	
Torque	Less Than 80.15
Pattern 3:	
Thrust force	Greater Than 183.25
Thrust force	Less Than 230
Pattern 4:	
Thrust force	Greater Than 158
Thrust force	Less Than 168.85
Positive patterns for the hole circularity at entry	
Pattern 1:	
Thrust force	Greater Than 40
Cutting force	Less Than 67.7
Pattern 2:	
Thrust force	Greater Than 150.5
Thrust force	Less Than 179.1
Pattern 3:	
Torque	Greater Than 705.3
Pattern 4:	
Cutting force	Greater Than103.85
Cutting force	Less Than 115
Negative patterns for hole circularity at entry	
Pattern 1:	
Thrust force	Greater Than 179.1
Cutting force	Less Than 103.85
Pattern 2:	
Torque	Less Than 80.15
Pattern 3:	
Cutting force	Greater Than 115
Torque	Less Than 705.3
Pattern 4:	
Thrust force	Greater Than121.6
Thrust force	Less Than 150.5
Positive patterns for hole diameter error at entry	
Pattern 1:	
Thrust force	Greater Than 52.15
Thrust force	Less Than 58.6
Negative patterns for hole diameter error at entry	
Pattern 1:	
Torque	Greater Than 137
Pattern 2:	
Thrust force	less Than 52.15
Positive patterns for hole diameter error at exit	
Pattern 1:	
Cutting force	Greater Than30.15
Cutting force	Less Than 41.45
Negative patterns for hole diameter error at exit	
Pattern 1:	
Torque	Greater Than 207.25
Pattern 2:	
Cutting force	less Than 30.15
Positive patterns for hole surface roughness	
Pattern 1:	
Thrust force	Greater Than 56.25
Thrust force	Less Than 58.6
Negative patterns for hole surface roughness	
Pattern 1:	
Torque	Greater Than 137
Pattern 2:	
Thrust force	less Than 56.25

In this chapter, detecting the thresholds values and characteristic patterns for machining conditions in term of uncontrollable variables, which are force and torque, using the LAD technique is presented. The key feature of the LAD is illustrated and its ability to detect characteristic patterns in the collected machining process data is shown. LAD technique is used in the diagnosis of machining outcomes by comparing each incoming new measurement to stored patterns. Because patterns are a conjunction of certain indicators' values, they can be used to build a decision boundary for classification by providing important information to distinguish observations which represent conforming products from those that are not. Experimental results are used as an illustrative example of the use of machine learning in order to monitor and detect the outcome of a machining process. This illustrative example was easy to visualize in 3D. In a more general case, when the number of indicators is greater than three, the generated patterns still give all the necessary and sufficient information in order to detect and to separate between the conditions that lead to conforming and nonconforming outcomes. LAD approach has the following advantages:

1. As it is shown in the application that is presented in this chapter, LAD produced high accuracy even if the data are non-separable. The generated patterns cover every observation at least once. There is no approximation in the generated patterns, and the generated patterns are represented in terms of the uncontrollable variables. This cannot be obtained when using the most traditional design of experiment technique or regression model.

2. LAD is not based on any statistical modelling. As such, any correlation between the uncontrollable variables, as well as the relation between the uncontrollable variables and the output, is reflected in the generated patterns. In contrast, although design of experiment can produce nonlinear models, isolated observations such as 4,5 and 31 in figure 3(c) will most probably be treated as outliers, and the results will be constrained by the obtained mathematical model.

3. LAD is dynamic; that is the attributes are monitored over time. This characteristic of LAD will be used in future work to detect any changes in the uncontrollable variables and to control the machining process on-line and in real time. LAD emphasizes the importance of keeping the historical data on the machining process. The larger the database is, the higher the chance of having a complete representation of the material's response to different machining conditions. For completely new material, design of experiments give a starting database from which analysis can be performed, but as the database increases in size, LAD and machine learning offer a more technologically advanced tool for detecting .

4. Over the past decades, quite complex modelling and experimentation was required to simulate the machining process. Modelling is based on assumptions and approximations. LAD is not limited by the search for a mathematical model to represent a complex machining process. It is

only limited by the computational capacity of a computer and the existence of data. Both conditions are in continuous development nowadays.

CHAPTER 3 ARTICLE 1: PROCESS CONTROL BASED ON PATTERN RECOGNITION FOR ROUTING CARBON FIBER REINFORCED POLYMER

Yasser Shaban, Mouhab Meshreki, Soumaya Yacout, Marek Balazinski, Helmi Attia

Published in:
Journal of intelligent manufacturing

(DOI: 10.1007/s10845-014-0968-6)

3.1 Abstract

Carbon Fiber Reinforced Polymer (CFRP) is an important composite material. It has many applications in aerospace and automotive fields. The little information available about the machining process of this material, specifically when routing process is considered, makes the process control quite difficult. In this paper, we propose a new process control technique and we apply it to the routing process for that important material. The measured machining conditions are used to evaluate the quality and the geometric profile of the machined part. The machining conditions, whether controllable or uncontrollable are used to control part accuracy and its quality. We present a pattern-based machine learning approach in order to detect the characteristic patterns, and use them to control the quality of a machined part at specific range. The approach is called Logical Analysis of Data (LAD). LAD finds the characteristic patterns which lead to conforming products and those that lead to nonconforming products. As an example, LAD is used for online control of a simulated routing process of CFRP. We introduce the LAD technique, we apply it to the high speed routing of woven carbon fiber reinforced epoxy, and we compare the accuracy of LAD to that of an Artificial Neural Network (ANN), since the latter is the most known machine learning technique. By using experimental results, we show how LAD is used to control the routing process by tuning autonomously the routing conditions. We conclude with a discussion of the potential use of LAD in manufacturing.

Keywords: Machining, Process Control, Logical Analysis of Data, CFRP, pattern recognition, knowledge extraction.

3.2 Introduction

The composite materials have special properties which make them the backbone of some industries such as aerospace, sporting, automotive and aircraft structure (Rahman et al. 1999). CFRP has very high modulus of elasticity, high tensile strength, low density, and high chemical stability. Most studies of CFRP are restricted to material properties and theoretical mechanics. Nowadays, the economic impact has an important consideration in manufacturing; therefore, it's important to study the machining process control for CFRP because it affects the production process (Ferreira et al. 1999). The machining of composite materials is more difficult than the machining of metals because they have non-homogeneous composition, and abrasive properties of reinforcing fibers. The cutting tool confronts fibers and matrix whose response to machining process could be completely different (Teti 2002). The complicated reaction of CFRP to machining, and consequently the defects which are introduced into the workpiece, in addition to the special required specifications of the machined part are the main reasons for the search of new techniques for process control.

Milling is used frequently in manufacturing in order to produce, with composite materials, parts which have high accuracy and high surface quality (Teti 2002), such as delamination, surface roughness and machined part dimensions (Davim and Reis 2005). Researches have been conducted on milling process control. They used artificial intelligence learning techniques in order to control machining process. For example, (Zuperl et al. 2012) used ANN and fuzzy logic to control the cutting force in the process of ball-end milling, and in order to maintain constant roughness. In (P. B. Huang 2014), The author developed an intelligent neural-fuzzy model for surface roughness monitoring system in milling operations. He developed a decision-making system which analyzed the cutting forces and then responded with an accurate output. He concluded that his developed system can be used, in future, as an adaptive control system of the machining parameters in smart Computer Numerical Control (CNC) machine. In (Zhang et al. 2007), the authors used ANN to develop surface roughness adaptive control in turning process. They used data from controllable cutting parameters such as feed rate, cutting speed, and depth of cut, and also uncontrollable monitored parameters such as vibration signals in order to develop neural-networks-based surface roughness adaptive control system. Other researchers used other techniques, for example, (Coker and Shin 1996) used ultrasonic sensing to control surface roughness during machining processes. (H. Wang and Huang 2006) used the concept of an Equivalent Fixture Error (EFE) to improve machining process control. Based on simulated data, they illustrated their concept. In (Du et al. 2012), the authors developed a robust approach for root causes identification in machining process using hybrid learning algorithm and engineering-driven rules. In order to judge whether the process is in normal or abnormal condition, off-line pattern match relationships and on-line time series measurements were used. They validated the developed approach by using data from the real-world cylinder head of engine machining processes. Due to the nonlinearity and complexity of milling process, traditional approaches fail to develop appropriate model to control the process (Haber et al. 2002). In (Landers et al. 2002), the authors concluded that the future of the milling process monitoring and control needs techniques that can determine threshold values and characteristic patterns which can be used to control and tune autonomously the controllable machine conditions (feed, cutting speed, etc.), on-line and off-line, in order to improve part accuracy.

In this paper, we present a pattern-based machine learning technique called Logical Analysis of Data (LAD). We use this technique in order to discover and to understand the hidden correlation between the machining variables of CFRP. Information is extracted from experimental results, and is presented in the form of characteristic patterns. These are hidden rules that characterize the temporal evolution of the machining process. Subsequently, these rules are used in machining process control. In section 3.3, the experimental procedure and results are presented. LAD approach is presented in section 3.4 and a numerical example is introduced. In section 3.5, the learning process, from the obtained experimental data, is introduced and comparison between LAD and the ANN is presented. In section 3.6, a simulated machining process control is used for building online-decision making procedure using LAD. Concluding remarks are given in section 3.7.

3.3 Experiment Description

The composition of the tested CFRP composite is quasi-isotropic laminate comprising 35 plies of 8-harness satin woven graphite epoxy prepreg with a final cured thickness of 6.35 ± 0.02 mm. The tool materials is 6.35 mm, four-flute, solid carbide end mill. The equipment is a Makino A88Ɛ machining center. In order to reach a spindle speed up to 40,000 rpm, an IBAG spindle speed attachment, which has a 1 kW power, is used. The routing tests is performed using four values of Spindle speed (rpm): 10,000, 20,000, 30,000, and 40,000, three values of feed (mm/min): 250, 500, and 1,000, and three values of tool overhang lengths (TL): TL1 =38 mm, TL2 = 31 mm, and TL3 = 24 mm. The experiments are repeated each 32 mm of cutting distance for three times. As such, we have three values of cutting distance (C): 32, 64, and 96 mm. In total, we have three feed rates (f), four cutting speed (v), three overhang length (TL), and three cutting distances (C); therefore, the total number of observations (experiments) is 108. This is a full factorial design of experiments. During slotting, the cutting forces are measured using a Kistler dynamometer 9255B, and the temperatures are measured using a FLIR ThermoVision A20M infra-red camera. For example, Trends for the feed force (Fx), transverse force (Fy), and axial force (Fz) for different speeds, feeds and tool overhang length (TL1 = 38mm) are shown in

Figure 3-1 The machined slots were characterized in terms of surface roughness, and delamination. The conforming specifications of these qualities are as follows:

- Exit and entry delamination ≤1%.
- Slot surface roughness right and left ≤ 1.2μm.

Schematic of the experimental setup is shown in Figure 3-2. A sample of the collected data are presented in Table 3.1. The observations that satisfy (don't satisfy) any of these specifications are identified by 1(0) in Table 3.1.

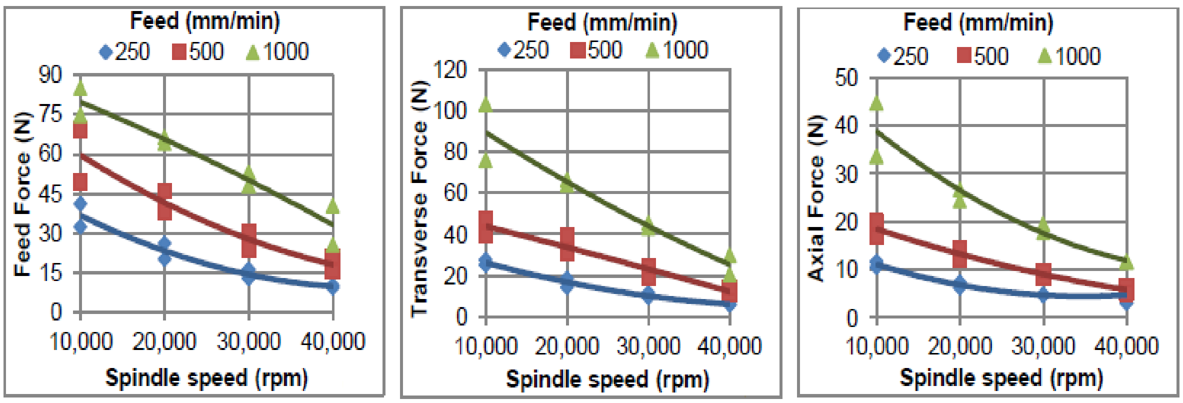


Figure 3-1: Trends for the feed force (Fx), transverse force (Fy) , and axial force (Fz) for different speeds, feeds and tool overhang length (TL1 = 38mm) (Meshreki et al. 2012).

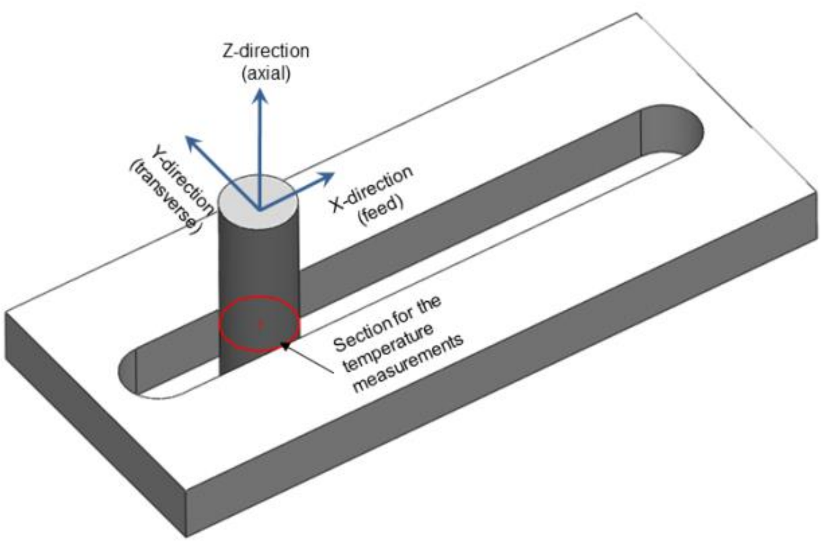


Figure 3-2: Schematic of the experimental setup

Table 3.1: A sample of the experimental results

Variables									Quality outcomes					
Controllable					Uncontrollable(monitored)									
No (1)	v rpm 10^4 (2)	f mm/min (3)	TL mm (4)	C mm (5)	Fx N (6)	Fy N (7)	Fz N (8)	Tmean ℃ (9)	Entry Delam. (10)	Exit Delam. (11)	ID (12)	Surface Rough. Right µm (13)	Surface Rough. Left µm (14)	ID (15)
1	4	250	38	32	9.2	5.8	6.5	305	0.068	0.0900	**0**	5.950	5.48	**0**
2	4	500	38	32	15.4	11.2	6.6	385	0.108	0.1009	**0**	6.520	6.33	**0**
3	4	1000	38	32	25.5	20.5	11.5	438	0.126	0.1301	**0**	8.020	6.67	**0**
4	3	250	38	32	12.8	9.4	4.9	319	0.122	0.1337	**0**	9.180	7.58	**0**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
25	4	250	24	32	18.4	5.7	3.2	305	0.001	0.001	**1**	0.85	1.03	**1**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
35	1	500	24	32	41.9	35.2	18.3	203	0.0016	0.0016	**1**	1.840	1.84	**0**
36	1	1000	24	32	68.9	97.4	45.5	301	0.053	0.0157	**0**	2.53	3.61	**0**
37	4	250	38	64	11.6	7.1	9.1	292	0.089	0.1009	**0**	5.80	5.32	**0**
38	4	500	38	64	19.4	11.9	7.6	387	0.213	0.1301	**0**	6.70	5.82	**0**
39	4	1000	38	64	36.2	25.9	12.2	452	0.414	0.1446	**0**	9.05	7.63	**0**

Table 3.2: A sample of the experimental results (continued)

40	3	250	38	64	15.6	10.9	5.0	319	0.2176	0.1409	**0**	8.580	5.92	**0**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
50	4	500	31	64	20.7	11.4	5.7	342	0.1046	0.0499	**0**	4.540	4.58	**0**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
67	2	250	24	64	24.2	11.2	6.1	220	0.0098	0.0098	**1**	1.160	1.07	**1**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
74	4	500	38	96	24.1	15.0	5.5	460	0.1009	0.1082	**0**	7.040	5.68	**0**
75	4	1000	38	96	42.0	27.6	12.2	519	0.2394	0.1373	**0**	7.910	6.44	**0**
76	3	250	38	96	15.8	10.7	4.8	389	0.0499	0.1155	**0**	8.260	6.97	**0**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
90	3	1000	31	96	42.3	31.3	15.3	441	0.1154	0.0754	**0**	7.340	4.89	**0**
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
100	3	250	24	96	23.2	8.0	4.9	281	0.0353	0.0170	**0**	1.140	1.02	**1**
:	:	:	:	:					:	:	:	:	:	:
107	1	500	24	96	57.9	41.3	21.5	371	0.0016	0.0134	**0**	1.560	1.84	**0**
108	1	1000	24	96	79.5	96.0	49.1	405	0.0317	0.0280	**0**	3.320	3.46	**0**

3.4 Logical analysis of data (LAD)

3.4.1 The methodology

LAD is a knowledge discovery approach that allows the classification of phenomena based on knowledge extraction and pattern recognition. It is applied in two consecutive phases, training or learning phase, where part of the database is used to extract special features or patterns of some phenomenon, and the testing or the theory formation phase, where the rest of the database is used to test the accuracy of previously learned knowledge. LAD uses a supervised learning technique; this means that the historical data or the database contains the variables and their corresponding outcomes or classes. For example, in Table 3.1, columns 2 to 9 are the variables, and columns 12 and 15 are the classes. In this paper, we use a two-class LAD technique. A multi-class LAD technique can be found in (M.-A. Mortada et al. 2011, 2013). After the two previously mentioned phases, new observations are introduced to LAD in order to be classified. This classification allows us to predict the quality outcome. The main advantages of LAD are: (1) LAD has explanatory power and causality identification which can be very useful in addressing machining process problems. This means that the user can track back any results, caused by a phenomenon or its effects, to its possible causes. This property appears particularly special, when LAD is compared to ANN which is characterized by the difficulty in determining the network structure and the number of nodes, and also the difficulty of interpreting the classification process. The ANN is a “black box” type of technique, which classifies new points without any explanations. (2) Unlike the statistical techniques which depend on distributions, and independence among variables, LAD is a non-statistical, non-approximate technique. LAD does not assume that the data belongs to any specific statistical distribution. (3) Unlike rules based on expert systems and expert knowledge, LAD extracts the knowledge hidden in the data. It then accumulates and preserves this knowledge which can be used at any time by the user, even if the human expertise is not available anymore. (4) No restriction concerning the type of data that LAD can deal with. LAD is capable of handling different types of data, whether nominal or numerical, discrete or continuous, simultaneously.

LAD was proposed for the first time at the Rutgers University Center for Operations Research (RUTCOR) (Peter L Hammer 1986) . The main steps of the LAD are the binarization of data, the pattern generation, and the theory formation. The objective of data binarization is the transformation of a database of any type into a Boolean database by using cut points technique. Many researchers presented different binarization techniques (Eddy Mayoraz and Moreira 1999). In this paper, we use the binarization technique that is presented in (Bores et al. 2000). The technique starts by ranking, in ascending order, all the distinct values, u , of a variable, then cut-points ε is inserted between each two values that belong to different classes. The cut-point is calculated as the average of the two values. A binary attribute is then formed from each cut-point such that:

$$b = \begin{cases} 1 & if \;\; u \geq \varepsilon \\ 0 & if \;\; u < \varepsilon \end{cases}$$

The number of transitions between distinct values from two different classes, and vice versa, is equal to the number of cut-points which leads to the total number of binary attributes replacing a numerical variable.

The objective of pattern generation is to find the characteristic patterns that differentiate between classes that are commonly called positive and negative. The positive (negative) class is a set, π^+ (π^-), of observations that belong to this class. Many techniques were proposed for pattern generation such as heuristics (Peter L Hammer 1986; P.L. Hammer and Bonates 2006), enumeration (Bores et al. 2000),column generation (Hansen and Meyer 2011) ,and linear programming(Ryoo and Jang 2009). In this paper, we follow the pattern generation technique which is proposed in (Ryoo and Jang 2009). The authors converted the pattern generation problem to a set covering problem, and solved it by a mixed integer linear programming (MILP) without any assumptions. Each positive observation $i \in \pi^+$ is represented as a Boolean observation vector $a_i = (a_{i,1}, \ldots a_{i,q}, a_{i,q+1} \ldots a_{i,2q})$. Each generated pattern p is associated with a Boolean pattern vector $W = (w_1, w_2, \ldots w_q, w_{q+1}, w_{q+2}, \ldots w_{2q})$ with size n , where n= 2q, q is the size of a binary observation vector.

A literal is a Boolean variable x or its negation $\bar{x}$ (Bores et al. 2000). A pattern p cannot include both the literal x_j and its negation $\bar{x}_j$ at the same time, thus the constraint $w_j + w_{j+q} \leq 1 \quad \forall\, j = 1,2, \ldots q$ must be respected. The number of literals used to define the pattern is called the degree of a pattern d. Pattern p of degree d is a conjunction of d literals; therefore, the pattern p is found after getting the Boolean pattern vector W which is the solution of the set-covering problem. For the generation of a positive pattern p+, that is a pattern that covers observations which belong to the positive class, $Y = (y_1, y_2, \dots y_{D^+})$) is the Boolean coverage vector whose number of elements equal to the number of positive observation D^+, and where y_i is equal to 0 if a pattern p+ covers a positive observation i, and 1 otherwise. Minimizing Y means finding a positive pattern that covers the maximum number of observations of this class. Our objective is to find a pattern that covers a maximum number of positive observations. This pattern is subsequently used to characterize the positive class. It is an indication of the unknown outcome or class. In this optimization problem, the decision variables are the pattern vector W, the degree d, and the coverage vector Y. By definition, a positive pattern cannot cover any negative observations, so the dot product of the pattern vector W and the observation $i \in \Pi^-$ must be less than the degree d of the pattern p, and for that reason the constraint $\sum_{j=1}^{2q} a_{i,j} w_j \leq d - 1 \;\; \forall\, i \in \Pi^-$ must be satisfied. Since the generated pattern doesn't have to cover all the observations in π^+ , the following constraint must be satisfied, $\sum_{j=1}^{2q} a_{i,j} w_j + q y_i \geq d \;\; \forall\, i \in \Pi^+$. The set covering problem is repeated until all observations in one class are covered by a set of generated patterns such that each observation is covered by at least one pattern. In order to speed-up the pattern generation procedure, the newly-generated pattern must not be a subset of the set of patterns that have already been generated. Every generated pattern vector W is stored as vector v in the set V containing all pattern vectors of the patterns generated previously. This condition can be formulated as:

$$\sum_{j=1}^{2q} v_{k,j} w_j \leq d_k - 1 \quad \forall\, v_k \in V.$$

In addition to the previously mentioned constraints, the problem can be summarized as follows:

$$\min_{W,Y,d} \sum_{i \in \Pi^+} y_i$$

$$\text{s.t.} \begin{cases} \sum_{j=1}^{2q} w_j = d \\ 1 \leq d \leq q \\ W \in \{0,1\}^{2q} \\ Y \in \{0,1\}^{D^+}. \end{cases} \tag{1}$$

After generation of the strongest pattern, which is the pattern that covers a maximum number of observations in the positive class, looping mechanism is used in order to generate an entire set of patterns that cover all the positive observations at least once. The same process is then repeated to obtain the negative patterns by using the set π^- of negative observations. A theory is then formed and a decision model is obtained.

The theory formation is the final step in LAD. A discriminant function, such as the one given in equation (2), is formulated in order to calculate a score ranging between -1 and 1. When the output of a discriminant function has a positive (negative) value, this means that the tested observation belongs to the positive (negative) class. Zero value means the evidences are not enough in order to decide to which class an observation belongs (M.-A. Mortada et al. 2011).

$$\Delta(O) = \sum_{i=1}^{N^+} \alpha_i^+ P_i^+(O) - \sum_{i=1}^{N^-} \alpha_i^- P_i^-(O) \tag{2}$$

Where $N^+(N^-)$ is the number of positive (negative) generated patterns, $P_i^+(O)\left(P_i^-(O)\right)$ is equal to 1 if pattern (i) covers observation O, and is equal to zero otherwise, $\alpha_i^+(\alpha_i^-)$ are the weights of the positive (negative) pattern $p_i^+(p_i^-)$. These weights are the proportion of observations covered by each pattern. They represent the power of each pattern. A strong pattern is the most powerful and cover the highest number of observations.

3.4.2 Numerical Example

In order to explain the LAD methodology, we introduce the following numerical example. We assume that we have the following seven observations, and the corresponding measured qualities such as the surface roughness or the delamination. The quality takes a label or class 1 (0) to represent conforming (non-conforming) specification. We assume that the machining conditions measurements' are already changed to binary attributes b1 to b5, by using the procedure presented in section 3.1. We search for the combination of machining conditions, that are the characteristic patterns, which differentiate between parts which are conforming or non-conforming to specifications. The seven observations are shown in Table 3.2 in columns 2 to 6. Each observation i=1 to 7 is associated with a Boolean observation vector $a_i = (a_{i,1}, a_{i,2}, \dots a_{i,q}, a_{i,q+1}, a_{i,q+2}, \dots a_{i,2q})$, where q= 10. These are the literals of the observations, as in columns 2 to 6, and their negations. The Boolean observation vectors are shown in Table 3.2.

Table 3.3: collected Boolean observation vectors and their classes

No	b_1	b_2	b_3	b_4	b_5	a_i	$a_{i,1}$	$a_{i,2}$	$a_{i,3}$	$a_{i,4}$	$a_{i,5}$	$a_{i,6}$	$a_{i,7}$	$a_{i,8}$	$a_{i,9}$	$a_{i,10}$	class
1	0	1	0	1	1	1	0	1	0	1	1	1	0	1	0	0	1
2	1	1	0	1	0	2	1	1	0	1	0	0	0	1	0	1	1
3	1	1	1	0	0	3	1	1	1	0	0	0	0	0	1	1	1
4	1	0	1	0	0	1	1	0	1	0	0	0	1	0	1	1	0
5	0	0	0	1	1	2	0	0	0	1	1	1	1	1	0	0	0
6	1	1	0	1	1	3	1	1	0	1	1	0	0	1	0	0	0
7	0	0	1	0	0	4	0	0	1	0	0	1	1	0	1	1	0

$Y = (y_1, y_2, y_3)$ is the Boolean vector whose number of elements equal to the number of positive observations, and where y_i is equal to 0 if a pattern p covers the positive observation i, and 1 otherwise. Minimizing Y means finding a positive pattern that covers the maximum number of positive observations, that is the strongest pattern.

Accordingly, the MILP for the pattern generation problem is formulated as follows:

Minimum $\quad y_1 + y_2 + y_3$

S.t.

$w_1 + w_6 \le 1, w_2 + w_7 \le 1, w_3 + w_8 \le 1, w_4 + w_9 \le 1, w_5 + w_{10} \le 1$

$w_2 + w_4 + w_5 + w_6 + w_8 + 5y_1 \ge d, w_1 + w_2 + w_4 + w_8 + w_{10} + 5y_2 \ge d$

$w_1 + w_2 + w_3 + w_9 + w_{10} + 5y_3 \ge d, w_1 + w_3 + w_7 + w_9 + w_{10} \le d - 1$

$w_4 + w_5 + w_6 + w_7 + w_8 \le d - 1, w_1 + w_2 + w_4 + w_5 + w_8 \le d - 1$

$w_5 + w_6 + w_7 + w_9 + w_{10} \le d - 1, w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_{10} = d$

$1 \le d \le 5, w_j \in \{0,1\}\ \forall\ j = 1, \dots, 10,\ y_1, y_2, y_3 \in \{0,1\},$

This MILP problem has three decision set of variables (y, d, w) and it can be solved by any MILP-solver (Linderoth and Lodi 2011). The strongest pattern was obtained as $W = (0,1,0,0,0,0,0,0,0,1)$ which means that $p_1^+ = x_2\bar{x}_5$, and therefore the attributes' values must be equal to (1, 0) at attributes(b_2, b_5) in order to be covered by this pattern. The pattern is of degree $d = 2$, Y= $(1,0,0)$ which means that there is one positive observation $(y_1 = 1)$ that is not covered yet. In this small example, it is easy to see that from the three positive observations 1, 2, and 3, observations 2 and 3 are covered by the pattern that is found, while observation 1 is not. The process of pattern generation is repeated in order to find a pattern that covers observation 1.

In order to generate the p_2^+ pattern, the observations which have been covered by p_1^+ are removed. The remaining data set is given in Table 3.3

Table 3.4: The remaining dataset after founding the first positive pattern

No	b_1	b_2	b_3	b_4	b_5	a_i	$a_{i,1}$	$a_{i,2}$	$a_{i,3}$	$a_{i,4}$	$a_{i,5}$	$a_{i,6}$	$a_{i,7}$	$a_{i,8}$	$a_{i,9}$	$a_{i,10}$	class
1	0	1	0	1	1	1	0	1	0	1	1	1	0	1	0	0	1
4	1	0	1	0	0	1	1	0	1	0	0	0	1	0	1	1	0
5	0	0	0	1	1	2	0	0	0	1	1	1	1	1	0	0	0
6	1	1	0	1	1	3	1	1	0	1	1	0	0	1	0	0	0
7	0	0	1	0	0	4	0	0	1	0	0	1	1	0	1	1	0

Let $Y = (y_1)$, where Y is the Boolean vector whose number of elements equal to the number of positive observation. The MILP is as follows:

Minimum $\quad y_1$

S.t.

$w_1 + w_6 \le 1, w_2 + w_7 \le 1,\ w_3 + w_8 \le 1,\ w_4 + w_9 \le 1,\ w_5 + w_{10} \le 1$

$w_2 + w_4 + w_5 + w_6 + w_8 + 5y_1 \ge d, w_1 + w_3 + w_7 + w_9 + w_{10} \le d - 1$

$w_4 + w_5 + w_6 + w_7 + w_8 \le d - 1, w_1 + w_2 + w_4 + w_5 + w_8 \le d - 1$

$w_5 + w_6 + w_7 + w_9 + w_{10} \le d - 1,\ w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 + w_9 + w_{10} = d$

$1 \le d \le 5, w_j \in \{0,1\}\ \forall\ j = 1, \dots, 10,\ y_1 \in \{0,1\},$

By solving the MILP for the second iteration, the strongest pattern is $W = (0,1,0,0,0,1,0,0,0,0)$ which means that $p_2^+ = \bar{x}_1 x_2,$ and therefore the attributes' values are (0, 1) at attributes(b_1, b_2).The pattern is of degree $d = 2$, Y= $[0]$ which means that all the positive observations are covered. Since all the positive observations are covered by at least one pattern, the pattern generation procedure is stopped. The same procedure is repeated in order to generate the negative patterns. Finally the generated patterns are:

Positive patterns:$p_1^+ = x_2\bar{x}_5$ with weight $\alpha_1^+ = 2/3$ and $p_2^+ = \bar{x}_1 x_2$ with weight $\alpha_2^+ = 1/3$

Negative patterns:$p_1^- = \bar{x}_2$ with weight $\alpha_1^- = 3/4$ and $p_2^- = x_1 x_5$ with weight $\alpha_2^- = 1/4$

The interpretability power of LAD is obvious from the fact that any user can now go back to the collected observations and check the existence of these patterns and their coverage, as well as their signs and their meanings. The hidden knowledge discovery property is also obvious, since even in this small example, a human mind will not discover these patterns easily. This pattern discovery process is done by using the software cbmLAD (c. Software 2012; Bennane and Yacout 2012) . It took less than 1 second. Finally we note that the MILP is a procedure for pattern generation and discovery only. This means that LAD does not suppose any mathematical modeling of any relation between the variables.

The discriminant function that generates a score ranging between -1 and 1 is as follow.

$$\Delta(O) = \sum_{i=1}^{N^+} \alpha_i^+ P_i^+(O) - \sum_{i=1}^{N^-} \alpha_i^- P_i^-(O) = \left(\frac{2}{3}p_1^+ + \frac{1}{3}p_2^+\right) - \left(\frac{3}{4}p_1^- + \frac{1}{4}p_2^-\right)$$

For example, for a new observation (1,0,0,1,0), the discriminant function is

$$\Delta(O) = \left(\frac{2}{3}(0) + \frac{1}{3}(0)\right) - \left(\frac{3}{4}(1) + \frac{1}{4}(0)\right) = -0.75$$

The classification decision for this new observation is predicted to be the negative class.

3.5 Peformance comparison

3.5.1 The ANN technique

ANN is the most famous and well known machine learning technique. It has high efficiency on adaptation and learning. For these reasons, it’s used widely as modeling tool in machining process (Benardos and Vosniakos 2002; Çaydaş and Ekici 2012). An ANN is generally composed of three types of layers: an input layer which accepts the input attributes and has the number of neurons equal to the number of attributes, hidden layers which have some number of neurons, and an output layer that has one neuron. The number of hidden layers and its neurons depend on the nonlinearity of the model .All neurons in any layer are interconnected to the neurons of the pre and after layers through weighted links (Sharma et al. 2008).

The input variables which are controllable and monitored uncontrollable, as well as the quality outcomes, that are the delamination and the surface roughness are shown in Figure 3-3 (A, B). We use four models. Model (A-1) has controllable variables as inputs, namely cutting speed, feed, tool overhang length, and cutting distance. The output is the delamination which can be conforming or non-conforming to specifications. Model (A-2) has controllable variables as inputs, namely the cutting speed, feed, tool overhang length, and cutting distance. The output is the surface roughness which can be conforming or non-conforming. Model (B-1) has the monitored uncontrollable variables, namely the forces in three coordinates and the mean temperature, as inputs. The output is the delamination which can be conforming or non-conforming. Model (B-2) has the monitored uncontrollable variables, namely the forces in three coordinates and mean temperature, as inputs. The output is the surface roughness which can be conforming or non-conforming.

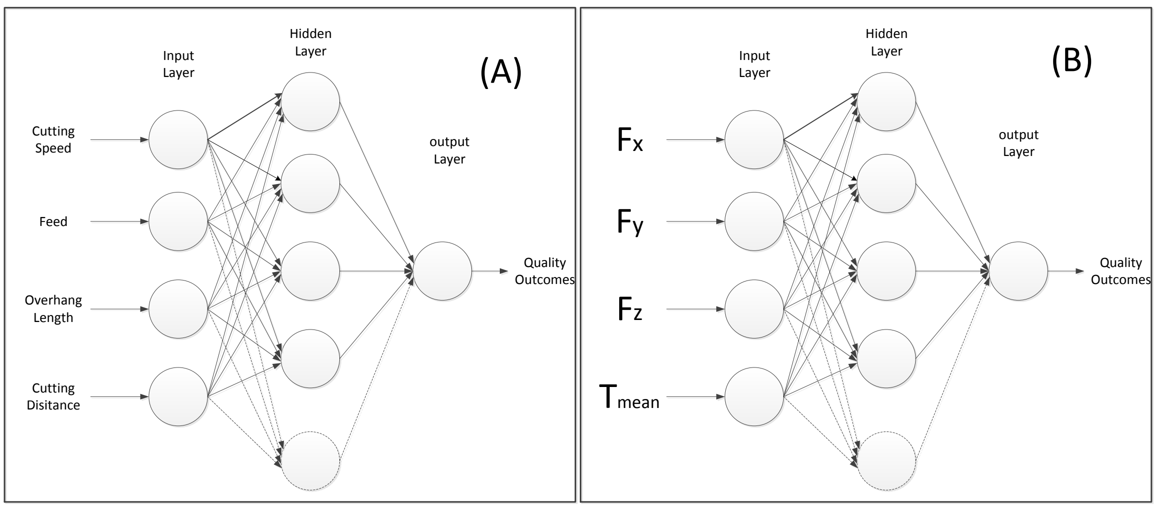


Figure 3-3: ANN models: (A) controllable variables model, (B) monitored uncontrollable variables model

Unlike the LAD approach, the ANN is subjected to the overfitting phenomenon. In order to find a good model, we tried several ones and we retained the best (Russell et al. 1995). By the best, we mean that we choose the network architecture that gives the highest prediction accuracy in the validation test. This will be discussed in details in section 5. In this paper, we use the Weka data mining software (Hall et al. 2009). For the delamination analysis, the proportion of conforming observation to non-conforming is 12 to 96 which is, obviously, an unbalance between minority and majority classes. The Synthetic Minority Over-sampling Technique (SMOT) is applied to rebalance and alter the class distribution(I.H. Witten et al. 2011). SMOT adjusted the relative frequency between two classes in the data to 48 to 96. The same technique is applied for surface roughness analysis to adjust the relative frequency from 6 to102 to 24 to102. For further reading about SMOT, we refer the reader to (Chawla et al. 2011). Table 3.4 shows the best obtained networks architecture.

The learning rate parameter takes a value between [0,1] in order to determine the step size, and hence how quickly the search converges. If it is either too large or too small, the search may overshoot and miss the minimum entirely, or slow the progress toward convergence. A momentum parameter term takes a value between [0,1]. It's used to update the value of a new weight by small proportion which leads to smooth searching process. The confusion matrix is $\begin{pmatrix} G & D^+ - G \\ D^- - H & H \end{pmatrix}$, where G is the total number of correctly classified positive observations, H is the total number of correctly classified negative observations, and $D^+(D^-)$ is the number of positive (negative) observations.

Table 3.5: ANN architectures for the four models.

The output	Controllable variables modelling	Monitoring variables modelling
DELAMINATION	***MODEL (A-1)*** (hidden layers No, its neurons)=(1,7) (Learning rate, momentum)=(0.3,0.2) Confusion matrix=$\begin{pmatrix} 44 & 4 \\ 7 & 89 \end{pmatrix}$	***MODEL (B-1)*** (hidden layers No, its neurons)=(1,5) (Learning rate, momentum)=(0.3,0.2) Confusion matrix=$\begin{pmatrix} 38 & 10 \\ 20 & 76 \end{pmatrix}$
SURFACE ROUGHNESS	***MODEL (A-2)*** (hidden layers No, its neurons)=(1,7) (Learning rate, momentum)=(0.3,0.2) Confusion matrix=$\begin{pmatrix} 21 & 3 \\ 3 & 99 \end{pmatrix}$	***MODEL (B-2)*** (hidden layers No, its neurons)=(1,3) (Learning rate, momentum)=(0.4,0.1) Confusion matrix=$\begin{pmatrix} 12 & 12 \\ 10 & 92 \end{pmatrix}$

3.5.2 Validation and comparison

The validation and the comparison between different techniques often represent a challenge for machine learning researchers (Wolpert 1996). Usually, two different learning techniques used for the same problem, and their results, are compared in order to decide which technique is better to use. By calculating the accuracy, which is obtained from cross-validation with several repetitions, the technique that has the higher accuracy is retained. This procedure is quite sufficient for comparison in many practical applications (I.H. Witten et al. 2011). In (P.L. Hammer and Bonates 2006; Soumaya Yacout 2010), LAD methodology was compared to the best reported results obtained by machine learning technique. The comparison was performed on a number of well-known problems which are conceived and kept in repositories in order to be used by researchers. The comparison was favorable to LAD technique (M. A. Mortada et al. 2009). In this paper, two qualities, namely the delamination and the surface roughness, are considered. For each one of the two qualities, the given specifications divide the outcomes space into two distinct spaces, the space of conforming products (positive) and the space of nonconforming products (negative). We also divide the set O of the n observations into two sets of training, L, and testing, T. In this paper we present the results obtained when the training set is composed of (n-1) observations and the testing is formed of the remaining observation. To calculate the classification accuracy we repeated the training-testing process n times, where each observation was chosen exactly once to constitute the testing set. This training and testing procedure is known as leave-one-out (LOOC) cross validation procedure, which is considered by many machine learning references as the best validation procedure when the amount of data for training and testing is limited (Ian H Witten and Frank 2011). LOOC is a special case of K-fold cross validation, where (K=n), n is the total number of observations. This procedure is attractive for two reasons. First, the greatest possible amount of data is used for training, which presumably increases the chance that the classifier is an accurate one. Second, the procedure is deterministic, which means no random sampling is performed. For example, if we divide the training set to equal parts, 50% for learning and 50% for testing, we omit 50 % of limited number of observations from the learning process, which affects negatively this process. Moreover, we will need a sampling strategy in order to choose 50% of the observations. In this paper, we present the results obtained when the training set is composed of (n-1) observations, and the testing is formed on the remaining observation. The procedure is then repeated n times. Two measures of accuracy, ACCURACY and the quality of classification, (ν) are used, where

$$\nu = \frac{a+b}{2} + \frac{e+g}{4}$$

The values (a) and (b) represent the proportion of observations, positive and negative, which are correctly classified. The values (e) and (g) represent the proportion of observations, positive and negative, which are not classified. Another measure is the ACCURACY = $\frac{G+H}{Nt}$, where G is the total number of correctly classified positive observations, H is the total number of correctly classified negative observations, and Nt is the total number of observations in the testing set. Table 3.5 shows the accuracy of the four models and the comparison between the accuracy of the ANN and the LAD techniques.

Table 3.6: Accuracy of the ANN and the LAD techniques.

Model	Quality outcome	ID	ANN		LAD	
			ACCURACY	ν	ACCURACY	ν
Controllable variables	Delamination	A-1	92.36	92.21	95.45	96.2
	Surface roughness	A-2	0.95	0.92	96.1	94.3
Monitoring variables	Delamination	B-1	0.79	0.79	81.24	80.02
	Surface roughness	B-2	0.70	0.83	86.2	88.22

In general, all statistical models are biased in one way or another; therefore, the comparisons between learning algorithms that are using different priors is meaningless (Wolpert 1996). Here, we compare between two different techniques, the ANN and the LAD. LAD methodology was compared to the most popular techniques of machine learning, such as Support Vector Machine (SVM), and ANN (P.L. Hammer and Bonates 2006; Soumaya Yacout 2010). In general, if the comparison shows that one of the algorithms has substantially high accuracy in comparison to the other, that algorithm should be used (Wolpert 1996). Obviously, it can be seen that the accuracy of LAD compares favorably with that of ANN.

3.6 Process control system

Our objective is to use the data presented in Table 3.1 in order to train LAD to detect automatically and without human interference, the threshold values and characteristic patterns for zones of machining conditions, that lead to acceptable quality, and those that lead to unacceptable quality. Although LAD generates positive and negative patterns for each of the four problems, in the following machining process control we use only the positive patterns of Models (A-1) or (A-2), and only the negative patterns generated for Models (B-1) or (B-2). In order to reach this objective, the software cbmLAD (c. Software 2012) is trained by using the data obtained from the experimental results that are shown in Table 3.1. Table 3.6 shows the positive characteristic patterns for Models (A-1) and (A-2), and the negative characteristic patterns for Models (B-1) and (B-2), which are found by the software.

Table 3.7: The positive patterns obtained by LAD for problems (A-1) and (A-2), and the negative patterns obtained for problems (B-1) and (B-2).

Positive patterns with controllable models	Pattern No	v rpm 10^4	f mm/min	Overhang length mm	cutting distance mm
Model (A-1)	1	<1.5	<375	<34.5	<48
	2	>3.5		<27.5	<48
	3	<2.5	>375,<750	<27.5	<48
	4	>1.5,<2.5	<375	<27.5	>48
	5	>1.5,<3.5	>375,<750	<27.5	>80
	6	<2.5	<375	<27.5	>80
Model (A-2)	1	>1.5	<375	<27.5	<48
	2	>2.5,<3.5	<375	<27.5	>80
	3	>1.5,<2.5	<375	<27.5	<80
	4	>1.5,<2.5	<750	<27.5	<48

Negative patterns with uncontrollable models	Pattern No	Fx N	Fy N	Fz N	T_{mean} °C
Model (B-1)	1		>5.735	>11.465	>210.5
	2	<24.2	>5.735	<8.9	
	3	>28.95		<10.845	
	4	<28.24	>5.735	>6.16,<9.755	>268.5
	5	<41.62	>19.185	>10.935	
Model (B-2)	1	>24.745			
	2	<23.2	>5.735		>210.5
	3	>23.385,<24.505			>225.5
	4	<23.2		>3.43	>194.5
	5	<14.91			

These generated patterns are used in the machining process control. The generated positive patterns illustrate the threshold boundaries for the controllable conditions that will always lead to conforming (positive) parts. In our machining process control, the negative patterns that are formed with the uncontrollable variables are used to give an alarm indicating that the machining process is beginning to produce unacceptable products. For example, the generated negative patterns (1) for Model (B-2) is Fx >24.745. This means that as long as Fx is higher than 24.745 the machined part will be non-conforming to the required specification of surface roughness. The same can be said for the negative pattern (5), which is Fx <14.91. These two constraints together illustrate the boundaries for the zone of Fx which should be avoided during the machining process. As we have explained, cbmLAD identifies and characterizes these regions perfectly and by using the lowest possible number of variables. To avoid the zones which are defined by the negative patterns, a simulated adaptive control loop is developed as shown in Figure 3-4.The generated patterns are incorporated in the machining process control which is shown in Figure 3-4.

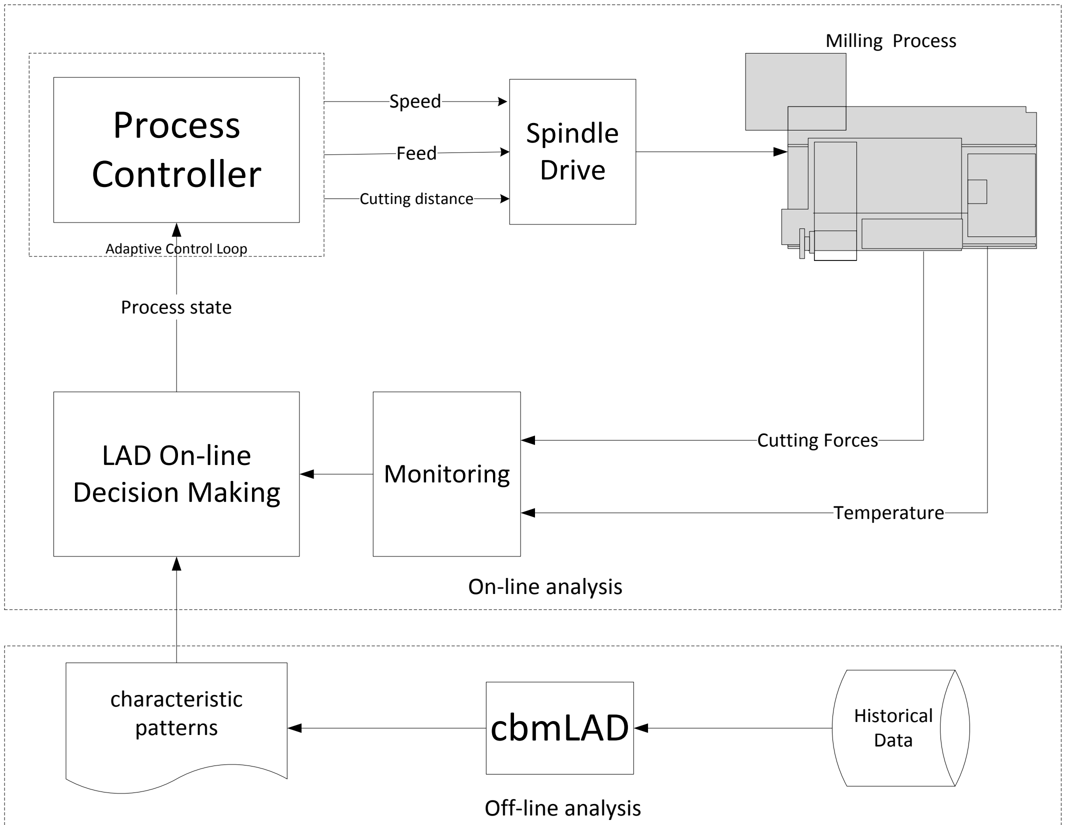


Figure 3-4: The machining process control

The machining process control is an adaptive control loop with an automatic adjustment of machining parameters, in our case the feed and speed, in order to improve operation productivity and part quality (Liang et al. 2004). Due to machine design constraints and complexity of finding monitoring parameters constraints and thresholds, process control loop is not commonly available in CNCs. Nevertheless, it attracts many researchers due to its potential to significantly improve operation productivity and part quality (Liang et al. 2004). In this paper, we assume that the machining process is monitored through sensors. Sensor's measurements are analyzed by the software cbmLAD in order to detect and identify the characteristic patterns; the patterns are obtained from the experimental data. They are then used in order to build the adaptive control loop. In Figure 3-4, a schematic diagram shows the machining process control. It starts by an off-line pattern generation by using cbmLAD. The generated patterns are transmitted to a "LAD On-line Decision Making" unit. The on-line loop starts by monitoring the uncontrollable variables. At each second, and by comparing the uncontrollable variables' values to the negative characteristic patterns, which are stored in "LAD On-line Decision Making" unit, a decision is made to whether change the values of the controllable values or to keep the current values. In the former case, the information is sent to the "Process Controller "unit in order to adjust the controllable variables to the nearest positive patterns' zones. The adjusted variables are the inputs to the actuator and the spindle drive.

In order to give a simulated example of the machining process control for the delamination quality, a simulated machining process control system is developed using labVIEW 8.5 software (Elliott et al. 2007). For example, we show in Figure 3-5 the front panels of Models (B-1) and (A-1). We use the negative patterns for the uncontrollable variables of Model (B-1), and the positive patterns for the controllable variables of Model (A-1), as shown in Table 3.6. The uncontrollable variables, which are the forces in three coordinates (Fx , Fy , Fz) and the mean temperature T_{mean}, are monitored and their values are sent to "LAD On-line Decision Making" unit every second in order to compare them to the stored negative patterns. A decision is then taken to either change the values of the controllable operating conditions in order to avoid the negative patterns' zones, or to keep them as they presently are. "LAD On-line Decision Making" gives an alarm if the uncontrollable variables comply with one of the negative patterns in Model (B-1). If the alarm is given, the "Process Controller" selects one of the positive patterns in Model (A-1). The selection of a positive pattern is guided by the dynamics of the machining process. The new values of the controllable variables are found in the selected positive patterns, and are the inputs to the actuator and the spindle drive. Adaptive control loop is looping at every second until "LAD On-line Decision Making" alarm is off. Figure 3-6 shows the flowchart for the process control.

In order to test the procedure that is described in the previous paragraph, a simulation model of the process control is developed. We assume that the correlation between controllable variable (speed (v), feed (f), tool overhang length (TL), and cut distance (C)) and the monitoring variables (forces in three coordinates (Fx, Fy, Fz) , mean temperature T_{mean}) for milling the CFRP composite material is represented by a simple multiple linear regression with a sample size (n) of 108. This assumption is only used in order to generate the values of the uncontrollable forces; in real life these values will be generated by the milling process itself, and they are captured by the sensors. The equations obtained using Weka data mining software were as follow:

$$\mathrm{Fx} = -0.0012\ v + 0.0402\ f + 0.4033\ \mathrm{TL} + 0.1913\ \mathrm{C} + 20.1253$$

$$\mathrm{Fy} = -0.0014\ v + 0.0493\ f + 0.5711\ \mathrm{TL} + 0.0719\ \mathrm{C} + 14.373$$

$$Fz = -0.0006\ v\ + 0.0235\, f\ +\ 14.9966$$

$$Tmean = 0.1952\ f\ +\ 7.4782\ TL\ +\ 1.0048\ C\ -\ 45.2817$$

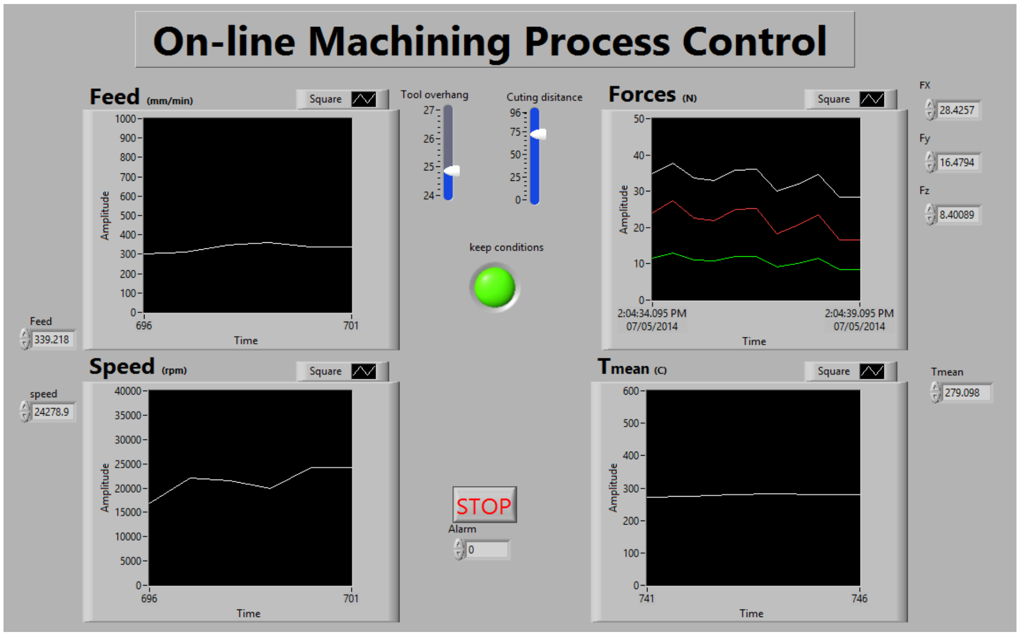


Figure 3-5: On-line machining process control for delamination quality using LabVIEW

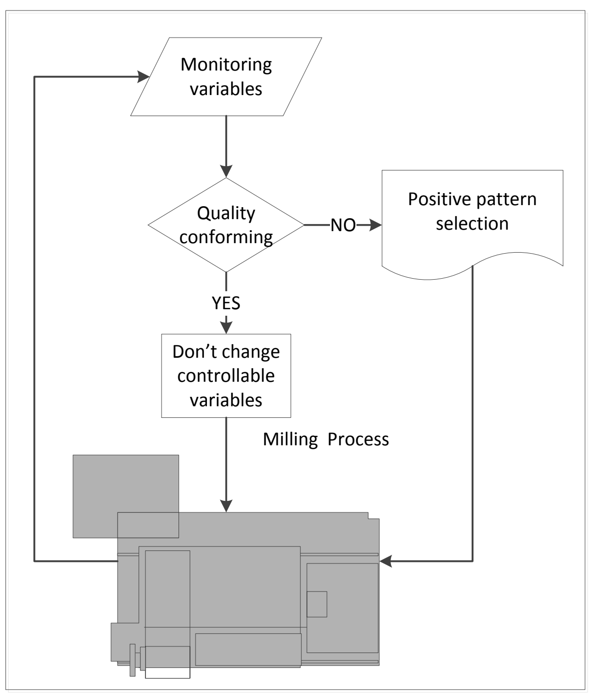


Figure 3-6: Flow chart of process control.

Since tool overhang length cannot be changed on-line, it was predefined and fixed before the simulation. According to the positive patterns 2 to 6 in Model (A-1), tool overhang length was restricted to less than 27.5. We chose an overhang length of $\mathrm{TL} = 24, 25, 26, \text{and} \;\; 27\ \mathrm{mm}$, for our simulated example. This means that only these five positive patterns of Model (A-1) are available to the "Process Controller" in order to control the machining process, since the first pattern can be satisfied with TL higher than 27 as long as it is less than 34.5. The cutting distance is also a predefined input which is set by the user before starting the simulation, and has a predefined value in the range of $\mathrm{C} \leq 96\ \mathrm{mm}$ during the simulation runs. For testing the simulated process control, we run the simulation model at $\mathrm{C} = 24, 27, 30, \dots, 87, 90, 93 \text{ and } \;\; 96\ \mathrm{mm}$ were performed. The total number of simulated runs are thus are equal to 100. As an example, Table 3.7 shows the results of how the iterations terminate by selecting one of the four positive patterns of Model (A-1). The elapsed time to find the positive pattern depends on the initial conditions, the inertia of CNC machine, and the number of positive patterns that were generated off-line, in this example we have four positive patterns. Run No 1 terminates after 13 seconds, by finding the positive pattern number (5) in Model (A-1), and run No 2 terminates in 4 seconds and found pattern number (4). In this work, we considered the iteration step as one second.

Table 3.8: Two runs of the simulated Process control using LabVIEW (continued)

Run No	Time sec	Controllable machining conditions				Uncontrollable (monitored)				**Pattern**
		Overhang length mm	cutting distance mm	v rpm	f mm/min	Fx N	Fy N	Fz N	$\mathbf{T_{mean}}$ ℃	
Run No 1	1	24	96	15000	450.00	48.25	36.16	16.57	318.40	Negative patterns
	2	24	96	26365	569.68	39.43	26.15	12.56	341.74	
	3	24	96	15916	641.84	54.87	44.34	20.53	355.81	
	4	24	96	23785	413.33	36.24	22.06	10.43	311.25	
	5	24	96	30121	731.22	41.41	28.86	14.10	373.24	

Table 3.8: Two runs of the simulated Process control using LabVIEW (continued)

Run No	Time sec	Controllable machining conditions				Uncontrollable (monitored)				Pattern
		Overhang length mm	cutting distance mm	v rpm	f mm/min	Fx N	Fy N	Fz N	T_{mean} °C	
	6	24	96	27314	603.44	39.64	26.49	12.78	348.32	
	7	24	96	21548	401.43	38.44	24.60	11.50	308.93	
	8	24	96	30710	517.75	32.13	17.51	8.73	331.61	
	9	24	96	19894	405.74	40.60	27.13	12.59	309.77	
	10	24	96	28787	554.72	35.92	22.02	10.76	338.82	
	11	24	96	33952	580.74	30.77	16.079	8.27	343.90	
	12	24	96	23776	482.74	39.04	25.49	12.07	324.79	
	13	24	96	32530	443.47	26.96	11.30	5.89	317.13	pattern 5 (positive)
	14	24	96	32530	443.47	26.96	11.30	5.89	317.13	
Run No 2	1	24	75	15000	450.00	44.22	34.65	16.57	297.18	Negative patterns
	2	24	75	21888	327.78	31.04	18.98	9.57	273.35	
	3	24	75	19682	335.56	34.00	22.45	11.07	274.87	
	4	24	75	24661	259.35	24.96	11.72	6.29	260.01	pattern 4 (positive)
	5	24	75	24661	259.35	24.96	11.72	6.29	260.01	

3.7 Conclusion

In this paper, LAD is applied to high speed routing of CFRP , and found the characteristic patterns that lead to conforming products and those which lead to nonconforming products, by exploiting the results obtained experimentally of a routing process of CFRP. LAD accuracy is compared to that of ANN. An on-line machining process control is developed by using the patterns that were found off-line. A simulated machining process control is implemented by using the experimental results, and LabVIEW software. The simulation model shows how LAD is used to control the routing process by tuning autonomously the routing conditions in order to always return to the machining zones defined by the positive patterns.

For the areas of further research, we are presently working on incorporating the machining process control in a real computer numerical control (CNC) machine. The learning phase will be done off-line by cbmLAD based on data obtained from sensors which are mounted to the CNC machine. At each unit of time, a new sensors' reading is transmitted to the unit "LAD On-line Decision Making". This latter works on-line in order to give and alarm each time a negative pattern of the uncontrollable variables is detected. The unit "Process Controller" searches on-line for a positive pattern of the controllable variables, then a decision to change the values of the controllable values or to keep the current values is taken. In the latter case, the actuator and the spindle execute the "Process Controller's" command. We are also working on studying the effects of initiating the alarm based on the discriminant function of the new observation instead of on only the appearance of a negative pattern.

CHAPTER 4 ARTICLE 2: OPTIMAL REPLACEMENT OF TOOL DURING TURNING TITANIUM METAL MATRIX COMPOSITES

Y. Shaban, M. Aramesh, S. Yacout, M. Balazinski, H. Attia, H. Kishawy

Published in:

Proceedings of the 2014 Industrial and Systems Engineering Research Conference

“This paper has won the first place prize of the Society for Engineering and Management System and 2014 Industrial and System Engineering Research Conference (ISERC 2014)”

4.1 Abstract

In machining of composite materials, little research has been conducted in the area of optimal replacement time of the cutting tool in terms of cost and availability. Due to the fact that tool failure represents about 20% of machine down-time, and due to the high cost of machining, optimization of tool replacement time is thus fundamental. Finding the optimal replacement time has also positive impact on product quality in terms of dimensions, and surface finish.

In this paper, we are finding the tool replacement time when a tool is used under constant machining conditions, namely the cutting speed, the feed rate, and the depth of cut, during turning titanium metal matrix composites (TiMMCs). Despite being expensive, MMCs are a new generation of materials which have proven to be viable in various fields such as biomedical and aerospace industrial. Proportional Hazard Model (PHM) is used to model the tool's reliability and hazard functions using Exakt software. Experimental data are obtained and used to construct and validate the PHM model, which is then used in decision making. The results are discussed and show that finding the optimal replacement time of the cutting tool is valuable in saving cost of machining process and maximizing the tool availability.

Keywords

Metal matrix composites, cost optimization, availability optimization.

4.2 Introduction

The economic factor's impact on tool life in machining is considered very important (Klim et al. 1996). Many researches tried to improve tool life by several ways such as using variable feeds during machining process (M Balazinski and Mpako 2000; Lin and Shyu 2000). The cutting tool cost dominates high percentage of the total machining cost. The tool cost represents around 25 per cent of the total machining cost (Sakharov et al. 1990). For this reason, finding the time at which a tool should be replaced is thus fundamental. The objective is to choose an optimal replacement time which results in low cost and high availability. If the tool is replaced earlier or later than necessary, valuable resources will be lost or products may be scrapped (Tail et al. 2010). Moreover, the tool replacement policy is one of the important aspects of tool management. Suitable tool management policy is important to reduce overall production costs (Jeang 1998).

In (V Makis 1995), the author used a PHM with a time–dependent covariate considering tool wear to find the optimal tool replacement time. In (Klim et al. 1996), the authors presented the effect of feed variation on tool wear and tool life. They proposed a new method to improve cutting tool life in machining. In (Tail et al. 2010), the authors used a PHM to model the tool's reliability and hazard functions, The PHM offers a good model for data representation. The cutting speed is considered as the model's covariate. In (Mazzuchi and Soyer 1989), the authors presented a PHM not only for modelling tool life, but also for evaluating the mechanisms attributed to the cause of tool failure. In (Ding and He 2011), they used a PHM for modelling the cutting tool wear reliability analysis. Vibration signals which are indication to tool wear are used as model's covariate. The PHM showed remarkable relationship between the tool condition monitoring information and the life distribution of tool wear. Many researchers consider the PHM as a good model for tool life. In most of these models, it was assumed that the tool wear has significant effect over the entire tool life. In this paper, the objective is to find the optimal replacement time which minimizes the cost and maximizes the availability during turning titanium metal matrix composites (TiMMCs). Ti-MMCs are a new generation of materials which have proven to be viable materials in various industrial fields such as biomedical and aerospace, and they are very expensive. The PHM is used to model the tool's reliability and hazard functions using Exakt software. The tool wear degradation is taken as model's covariate. In section 4.2, a brief description of the PHM is introduced, followed by the estimation of the model's parameters and the covariate's weight. In section 4.3, the optimal replacement policy for minimizing the cost and maximizing the availability is described. The decision rule which helps in decision making is introduced in section 4.4. In section 4.5, the experimental procedure which was carried out in order to collect data that is used for constructing the model is presented. The model developed and the final results are presented in section 4.6. Concluding remarks are given in section 4.7.

4.3 Model description

The PHM presents the failure rate as the product of a baseline failure rate $h_0(t)$, which is dependent only on the age (cutting time) of the tool, and a positive function ψ that represents the tool wear $Z(t)$. The failure rate at time t is thus expressed as in equation (1):

$$h\big(t, Z(t)\big) = h_0(t)\psi\big(Z(t)\big) \ \ for\ t \geq 0 \tag{1}$$

In this paper we consider a PHM with a baseline Weibull hazard function. The Weibull distribution is extensively used in modelling the time to failure due to its flexibility in modelling a variety of failure data. Using Weibull as a baseline function in modelling the tool failure was considered in (V Makis 1995; Tail et al. 2010; Mazzuchi and Soyer 1989) . This model is sometimes called the Weibull parametric regression model. It is given as follows:

$$h(t, Z(t); \beta, \eta, \gamma) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\{\textstyle\sum_1^m \gamma_i Z_i(t)\} \tag{2}$$

Where β is the shape parameter, η is scale parameter, m is the number of covariates which have effect on the hazard rate, and γ is the weight of each covariate. The covariates may be controllable variables such as cutting speed (v), feed rate (f), and depth of cut (a_p), or uncontrollable (monitored) variables such as the cutting forces, the tool wear, and the temperatures. In this paper all controllable covariates are kept constant, so they will not affect the analysis of the model. The wear is the only covariate which will be monitored at discrete points of time through inspections and the appropriate model is given in equation (3), where m=1,

$$h(t, Z(t); \beta, \eta, \gamma) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\{\gamma Z(t)\} \tag{3}$$

$Z(t)$ depicts the evolution of the covariate representing the wear which is monitored and measured at discrete intervals of time. It has a finite state space. In this paper we consider two sates; the normal and the failure states. This latter is defined by the tool maximum flank wear length (VB_{Bmax}) reaching a predefined level equal to 0.2 mm. The conditional survival function can thus be given as in equation (4),

$$R(t; Z) = P(T > t|Z(s), s \leq t) = exp\left(-\int_0^t h_0(s)\psi\big(Z(s)\big)ds\right)\ for\ t \geq 0, \tag{4}$$

Where T is the random variable that represents the time to failure of the tool. When using Weibull distribution, equation (4) is given as follows:

$$R(t; Z) = \ exp\{-\left(\frac{t}{\eta}\right)^{\beta} e^{\gamma Z(t)}\} \tag{5}$$

The conditional survival function $R(t; Z)$ and its derivative $\dot{R}(t; Z) = h\big(t, Z(t)\big)R(t; Z)$ are used to estimate the parameters (β, η, γ) by using maximum likelihood function (Banjevic et al. 2001).

4.4 Optimal replacement Policy

In 1978, Bergman (Bergman 1978) investigated the optimal replacement rule which is considered a control-limit value $(d > 0)$. He found that it is optimal to replace either at failure time T or at T_d, the preventive replacement time, when the state variable has reached some threshold, whichever occurs first. The optimal stopping rule is written in equation (6).

$$T_d^* = inf\{t \geq 0: Kh(t, Z(t)) \geq d^*\} \tag{6}$$

Where K is the difference between the failure replacement cost $C + K$ and the preventive replacement cost C. According to the theory of renewal reward processes, the expected cost per unit time can be expressed as:

$$\phi(T_d) = \frac{C\ P(T_d < T) + (C+K)\ P(T_d \geq T)}{W(d)} = \frac{C + K\ P(T_d \geq T)}{W(d)} \tag{7}$$

It has been shown that $d^* = \phi(T_d^*)$ is the optimal cost at which the $\phi(T_d)\ is\ minimum$ and T_d^* is the optimal time to replace. $P(T_d \geq T)$ is the probability of failure replacement, $P(T_d < T)$ is the probability of preventive replacement, and $W(d) = E(\min\{T_d, T\})$ is the expected replacement time. Optimal level d^* can be found by using the fixed-point iteration procedure (Banjevic et al. 2001; Viliam Makis and Jardine 1992) or by using Semi-Markovian Covariate Process (Bergman 1978). Similarly, we can represent the availability function as in equation (7).

$$A(T_d) = \frac{uptime}{uptime + downtime} = \frac{W(d)}{W(d) + T_p\ P(T_d < T) + (T_p + K)\ P(T_d \geq T)} \tag{8}$$

The optimal availability is achieved when $A(T_d)\ is\ maximum$, where T_d^* is the optimal time to replacement, T_p is the time required to perform the preventive replacement, and $T_f = (T_p + K)$ is the time required to perform failure replacement. We note that in equation (8), K is the difference between T_f, T_p, while in equation (6) it was the difference between the failure replacement cost and the preventive replacement cost.

4.5 Decision rule

The important question in tool replacement policy is "Should we keep running or should we replace the tool now?". The decision rule which can be derived from equation (6) gives the answer to this question, by monitoring the tool wear at discrete time intervals (Banjevic et al. 2001). From equation (6) we get:

$$Kh(t, Z(t)) \geq d^* \tag{9}$$

$$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma Z} \geq \frac{d^*}{K} \tag{10}$$

$$e^{\gamma Z} \geq \frac{d^*\, \eta^{\beta} t^{-(\beta-1)}}{K\beta} \tag{11}$$

$$\gamma Z \geq \ln\left(\frac{d^*\, \eta^{\beta}}{K\beta}\right) - (\beta-1)\ln t \tag{12}$$

$$Z^c(t) \geq \mathrm{g(t)} \tag{13}$$

The function g(t)= $ln\left(d^*\, \eta^{\beta}/K\beta\right) - (\beta-1)\ln t$ can be consider as "warning level" function, applied to an "overall" covariate value $Z^c(t)$.

4.6 Experiment description

Workpiece material: A cylindrical bar of Ti-6Al-4V alloy matrix reinforced with 10-12% volume fraction of TiC ceramic particles is used.

Tool material: TiSiN-TiAlN nano-laminate PVD coated grades (Seco TH1000 coated carbide grades) were utilized.

Equipment: We used a 6-axis Boehringer NG 200, CNC turning center in order to conduct experiments, as shown in figure (4-1).

Experimental details: Based on the recommendation of the tool supplier, the experiments have been conducted under the following constant cutting conditions: Cutting speed (v) =60 m/min, feed rate (f) =0.15 mm/rev, and depth of cut (a_p) =0.2 mm.

Figure 4-1: The experiment setup

Sequential inspections and turning tests are conducted for each tool in order to measure the wear. The wear is measured after each inspection by using an Olympus SZ-X12 microscope. The procedure continues until the tool wear threshold ($VB_{Bmax} = 0.2$ mm) is reached. The procedure is replicated for six tools. The collected data is shown in figure (4-2).

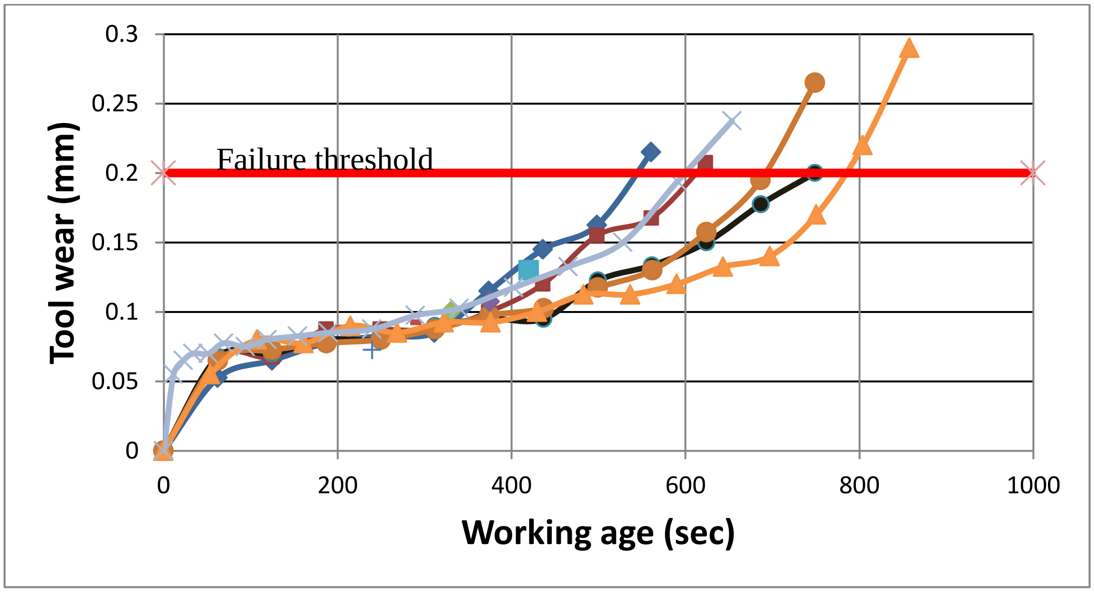


Figure 4-2: Tool wear measurements for 6 tools

In order to calculate the time to failure TTF, the wear evolution between two measurements (VB_i , VB_{i+1}) is assumed to be linear as in figure (4-3). The TTF is found at tool wear VB_{Bmax} = 0.2 mm by interpolating between (t_i , t_{i+1}). For example, from Table (4.1), and by interpolating between the fifteenth and the sixteenth inspections, then by using equation (14), the time to failure is found to be 782.73 sec. This interpolation is repeated for six tools. The results for the six tools and their inspections' results are given in Table (4.2). In this table, ID means the identification for tool from 1 to 6, B-event means the beginning for a new tool, IN-event means inspection process(measuring the wear), and EF-event means ending with failure (reaching the wear threshold).

$$\frac{\varepsilon}{\Delta t} = \frac{0.2 - VB_i}{\Delta VB} \quad , \; TTF = \varepsilon + t_i \tag{14}$$

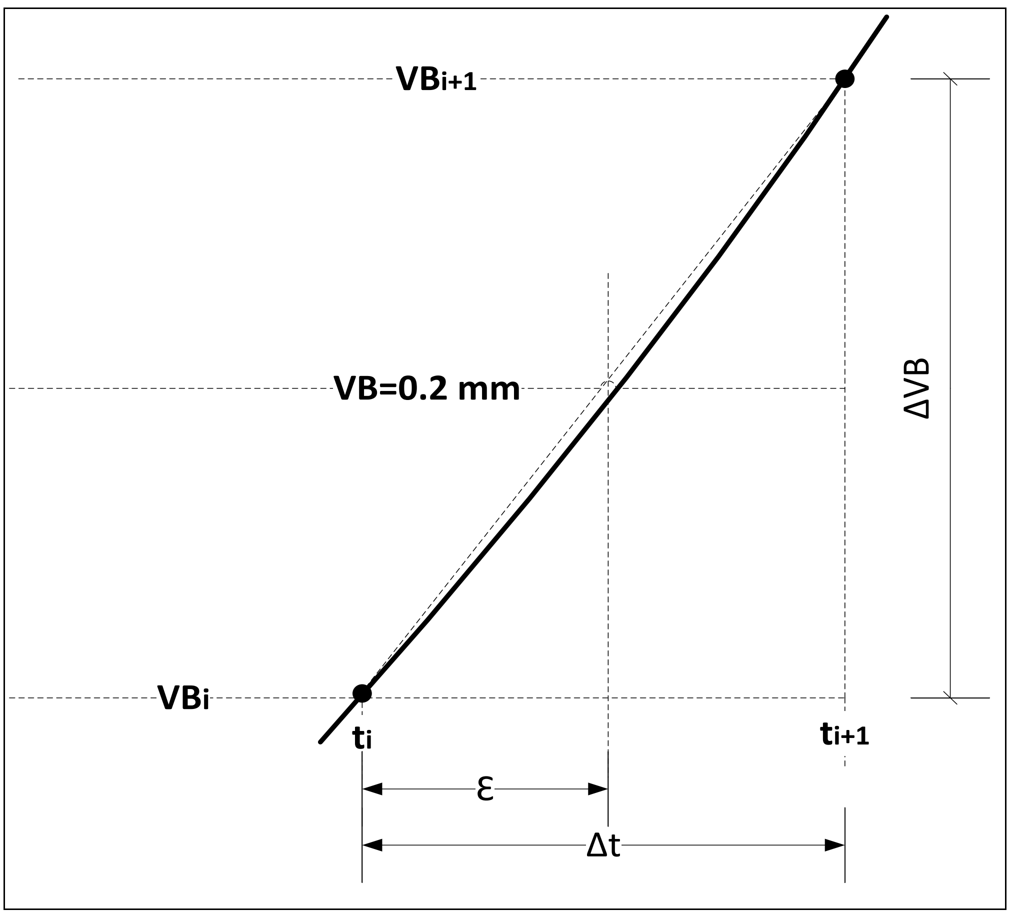


Figure 4-3: Wear interpolating

Table 4.1: The experimental results showing the wear of tool number 6

Inspection No	Time (sec)	VB_B (mm)
1	0	0
2	53.7	0.055
3	107.4	0.08
4	161.1	0.0775
5	214.8	0.09
6	268.6	0.085
7	322.3	0.0925
8	376	0.0925
9	429.5	0.1
10	483	0.1125
11	536.5	0.1125
12	590	0.12
13	643	0.1325
14	697	0.14
15	750	0.17
16	804.1	0.22

Table 4.2: Times to failure and samples of wear value inspections for the six tools

Tool ID	Working Age sec	Wear mm	Event	Tool ID	Working Age sec	Wear mm	Event
1	0	0	B	**4**	0	0	B
1	62.31	0.0525	IN	**4**	62.45	0.065	IN
1	:	:	IN	**4**	:	:	IN
1	:	:	IN	**4**	:	:	IN
1	:	:	IN	**4**	:	:	IN
1	498.47	0.1625	IN	**4**	686.69	0.195	IN
1	542.97	0.1625	EF	**4**	691.14	0.195	EF
2	0	0	B	**5**	0	0	B
2	11.29	0.055	IN	**5**	62.38	0.0675	IN
2	:	:	IN	**5**	:	:	IN
2	:	:	IN	**5**	:	:	IN
2	:	:	IN	**5**	:	:	IN
2	590.88	0.194	IN	**5**	686.72	0.1775	IN
2	599.54	0.194	EF	**5**	749.17	0.1775	EF
3	0	0	B	**6**	0	0	B
3	62.31	0.0675	IN	**6**	53.72	0.055	IN
3	:	:	IN	**6**	:	:	IN
3	:	:	IN	**6**	:	:	IN
3	:	:	IN	**6**	:	:	IN
3	560.93	0.1675	IN	**6**	750.63	0.17	IN
3	611.61	0.1675	EF	**6**	782.73	0.17	EF

4.7 Development the model and results

By using the software Exakt (Banjevic et al. 2001), The PHM parameters are estimated, and the resulting hazard function is given as follows in equation (15):

$$h(t, Z(t)) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma\, Z(t)} = \frac{3.713}{86330}\left(\frac{t}{86330}\right)^{2.713} e^{109.1\, Z(t)} \tag{15}$$

EXAKT offers Kolmogorov-Smirnov test (K-S test) to evaluate the model fit. The summary of goodness of fit test is automatically produced as in table (4.3).The test shows that the PHM offers a good modeling for the data.

Table 4.3: Summary of goodness of fit test results

Test	Observed value	P-value	PHM Fits Data
Kolmogorov-Smirnov	**0.378857**	**0.280343**	**Not rejected**

After determining The PHM the optimal replacement policy-cost analysis is performed. The optimal time to replacement T_d^* is calculated with a cost ratio of 2:1 (preventive replacement cost is estimated to be \$100, and the failure replacement cost is \$200, thus K is equal to \$100). As shown in figure (4-4), the cost value on the curve consists of the sum of the red portion that represents the unplanned failures cost, and the green portion which represents the preventive maintenance cost.

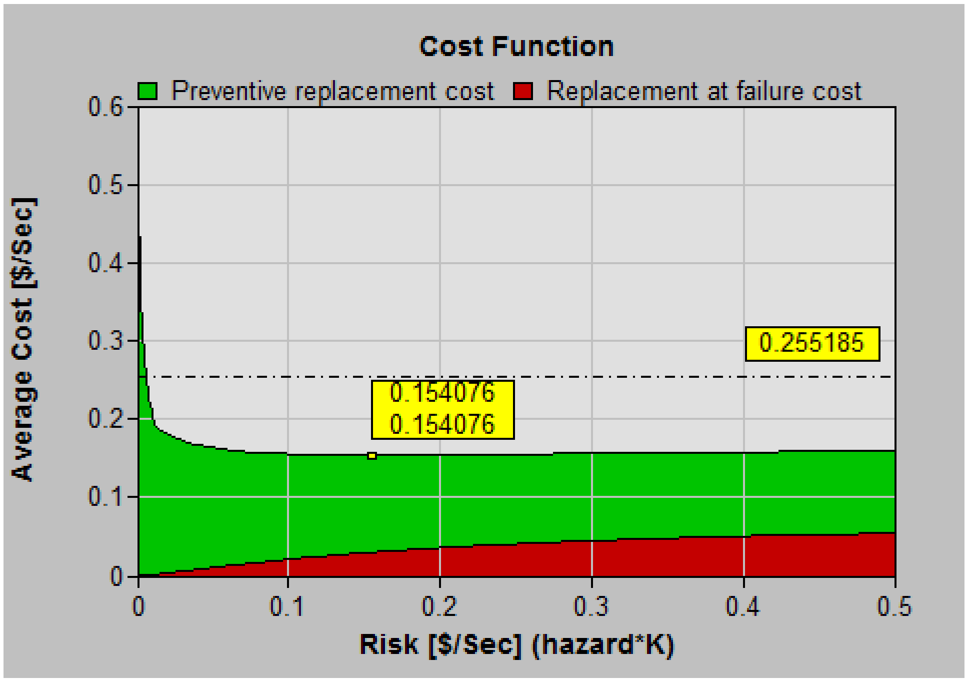


Figure 4-4: Condition-based replacement policy-cost analysis

Table (4.4) summarizes the information in figure (4-4). It compares the optimal cost $d^* = \phi(T_d^*)$ = 0.154 $, and and the expected time between replacements, 719 sec of the optimal policy, with those ($0.26 and 784 sec) of the "run to- failure" policy. It quantifies the expected preventive and failure costs ($0.124 and $0.03 respectively) in the optimal policy, and the percentage of incidences (89.2% will be preventive actions and 10.8% will be failure replacement action) achieved when the optimal policy is used. Finally, the table shows that the optimal policy proposes more interventions, on the average every 719 sec for the optimal policy versus 784 sec for the policy of ‘run- to- failure, in order to achieve a net per unit time saving (of $0.1 or 40%).

Table 4.4: Summary of cost analysis

	Cost [$/sec]	Preventive Repl.Cost [$/sec]	Failure Repl.Cost [$/sec]	Prev. Repl. [%]	Failure Repl. [%]	Expected time between Replacements
Optimal Policy	0.154076	0.124033 (80.5%)	0.0300427 (19.5%)	89.2	10.8	719.142
Replacement Only At Failure	0.255185	0 (0.0%)	0.255185 (100.0%)	0.0	100.0	783.746
Saving	0.101109 (39.6%)	-0.124033	0.225142	-89.2	89.2	-64.6044

Similarly, we found the optimal replacement policy that maximizes the availability. The optimal is conducted when the time required to preventive replacement, $T_p = 160$ sec, and the time required to failure replacement, $T_f = 540$ sec). From the results shown in figure (4-5) and table (4.5), it is found that the optimal availability $A(T_d^*\ )$is equal to 78.75%, and the and the expected time between replacements 843 sec for this optimal policy, while the availability and the time to replacement are equal to 59% and 784 sec, respectively, in the "run to- failure" policy. Practically speaking, we "buy" high availability by paying for it with more frequent interventions.

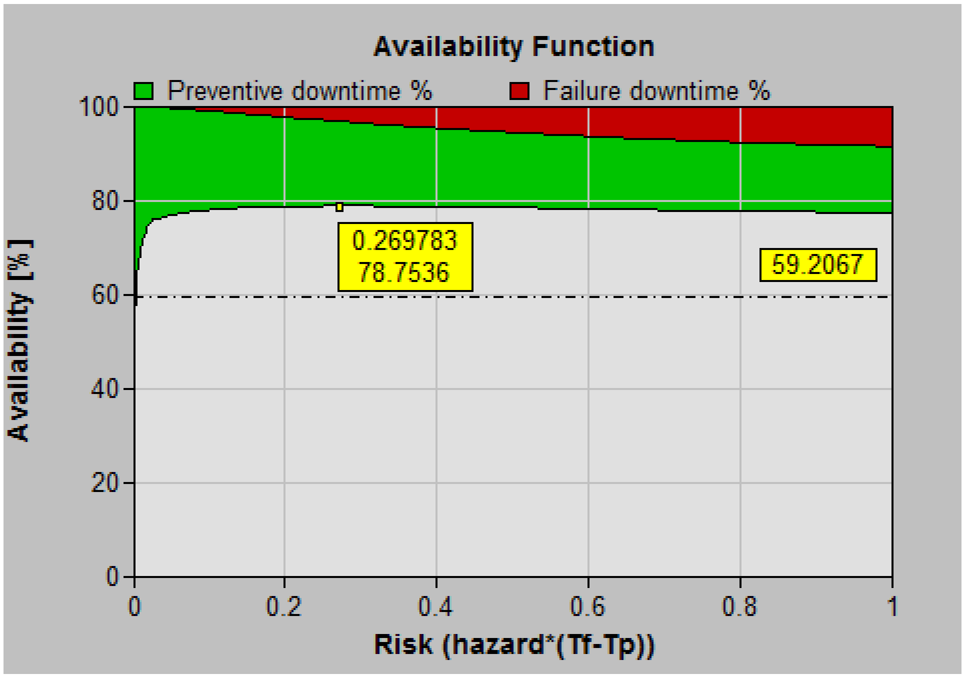


Figure 4-5: Condition-based replacement policy-availability analysis

Table 4.5: Summary of availabilty analysis

	Availability [%]	Preventive Downtime [%]	Failure Downtime [%]	Prev. Repl. [%]	Failure Repl. [%]	Expected time between Repl.[s]
Optimal Policy	78.75 (664.68*)	17.99 (84.69%)	3.25 (15.31%)	94.9	5.1	843.994 (179.318**)
Replacement Only At Failure	59.21 (783.75*)	0 (0.0%)	40.79 (100.0%)	0.0	100.0	1323.75 (540**)
Saving	(19.55%)	-17.99	37.54	-94.9	94.9	-479.752

* expected uptime, **expected downtime

In practice, the costs of failure ($C = 100\$, C + K = 200\$$), the planned inspection intervals (Δ= 60 sec) and the PHM model parameters are considered collectively in order to build the "warning level" function g(t)= $ln\left(d^* \eta^\beta / K\beta\right) - (\beta - 1)\, ln\, t$ as shown in figure (6). Once the decision model is built, we can make a decision that will optimize the long-run maintenance cost for the tool, or the long run availability of the machine. By defining the tool working age and the composite covariate $Z^c(t) = \gamma Z = 109.1 * wear$, the optimal decision is to determine whether the tool should be replaced immediately (the red area in figure (4-6)), or should we keep operating and be inspecting at the next inspection time (the green area), or should we keep operating but expect to replace before the next inspection time (the yellow area).

Moreover, the model was examined by using the data from previous histories to see what the decision model would have recommended for failed tool. The data in table (4.1) for tool (ID=6) is as shown in figure (4-6). According to equation (12), the decision chart gives us alert "intervene immediately" at working age 750.63 sec (inspection number 15 in table (4.1)) because the composite covariates $Z^c(t = 750.63) = \gamma Z = 109.1 * wear = 109.1 * 0.17 = 18.55$. This point crosses the "warning level" function g(t). Obviously, in this case, the model was capable of predicting the best action to make perfectly. The optimal replacement decision gives 'warning alert' before the tool's failure. We recapitulate the optimal decision policy in following words "the optimal policy suggests replacement at t for which $109.1 * wear \geq 34.41 - 2.713\, ln\, t$ ".

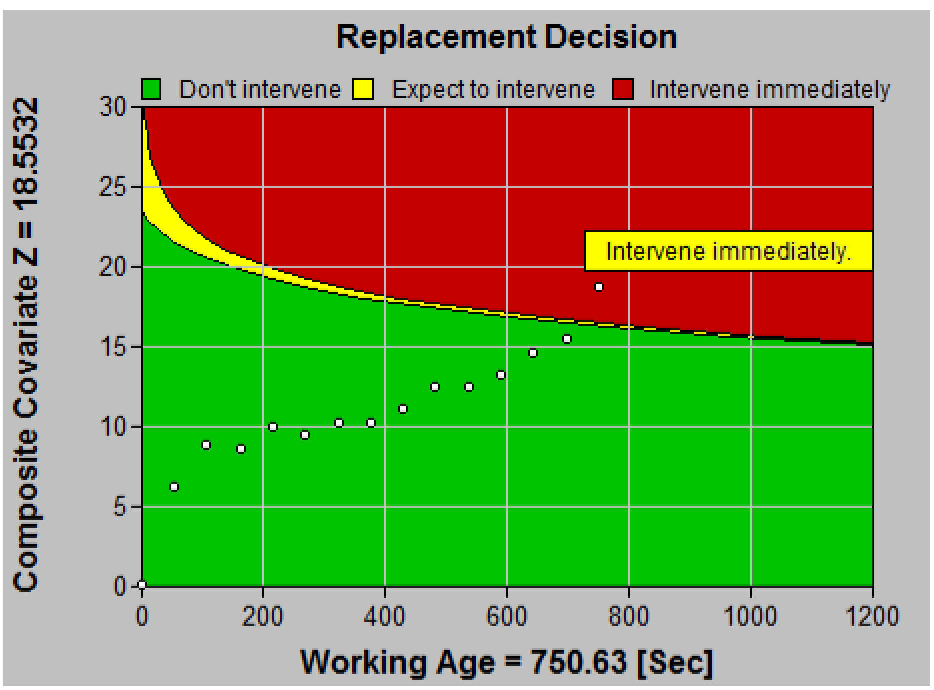


Figure 4-6: Condition-based replacement policy-optimal decision.

4.8 Conclusion

In this paper, experimental data were collected during turning titanium metal matrix composites (TiMMCs). The collected data were used to construct the PHM model which is then used to find optimal tool replacement time .The PHM offered a good modelling for the times to failure and tool wear degradation. The PHM models' parameters and economic objectives were considered to build the optimal decision chart. The study concluded that the optimal replacement times either lead to a cost reduction of 40 percent in case of cost analysis or lead to an increase of 79 percent in the case of availability analysis.

In future work, the tool wear will be monitored either directly by using a coupled device (CCD) camera which will follow the evolution of tool wear on-line, or indirectly by predicting the wear by monitoring the machining forces, and then by using a machine learning technique and then use it for decision making.

Identifying optimal replacement times for cutting tools

Yasser Shaban, Maryam Aramesh, Soumaya Yacout, Marek Balazinski, Helmi Attia, Hossam Kishawy

Submitted to:
Industrial Engineers (IE) magazine

"The first place winner of the SEMS Student Paper Competition, and recommended by SEMS for publication in the Industrial Engineers Magazine"

Overview of the article:

To answer the important question in cutting tool replacement strategy: “Should we keep running or should we replace the tool now?”

Introduction

The economic factor’s impact on tool life in machining is considered very important in tool management. Due to the fact that tool failure represents about 20 percent of machine down-time, and due to the high cost of machining, finding the time at which a tool should be replaced is thus fundamental. Many engineers tried to improve tool life by several ways such as using variable feeds during machining process. Their objective is to choose an optimal replacement time which results in low cost and high availability. If the tool is replaced earlier or later than necessary, valuable resources will be lost or products may be scrapped. Moreover, the tool replacement strategy is one of the important aspects of tool management. Suitable tool management strategy is important in order to reduce overall production costs.

In this work, we are finding the tool replacement time when a tool is used under constant machining conditions, namely the cutting speed, the feed rate, and the depth of cut during turning titanium metal matrix composites (TiMMCs). Despite being expensive, metal matrix composites are a new generation of materials which have proven to be viable in various fields such as biomedical and aerospace. Proportional Hazard Model (PHM) is used to model the tool’s reliability and hazard functions using EXAKT software. Experimental data are obtained and used to construct and validate the PHM model, which is then used in decision making. The results are discussed and show that finding the optimal replacement time of the cutting tool is valuable in saving cost of machining process and maximizing the tool availability.

Model description

The Proportional Hazard Model (PHM) presents the failure rate as the product of a baseline failure rate in the form of a Weibull hazard function which is dependent only on the age (cutting time) of the tool, and an exponential positive function that represents the conditions at which the process is functioning. These conditions are represented by covariates which may be controllable variables such as cutting speed, feed rate, and depth of cut, or uncontrollable (monitored) variables such as the cutting forces, the tool wear, and the temperatures. This means that the failure rate of the cutting tool is not only dependent on its age, but it is also affected by the covariates.

The Weibull distribution is extensively used in modelling the time to failure due to its flexibility in modelling a variety of failure data. This model is sometimes called the Weibull parametric regression model. We keep all controllable covariates constant, so they will not affect the analysis of the model. The wear is the only covariate that is monitored at discrete points of time through inspections. In this paper, we consider two states; the normal and the failure states. The latter is defined by the tool flank wear length reaching a predefined threshold equals 0.2 mm. We now turn to the discussion of the tool replacement strategies.

Optimal replacement Strategy

In 1978, Bo Bergman investigated the optimal replacement rule which is considered a control-limit value d^*. He found that it is optimal to replace a tool either at failure time or at the preventive replacement time, when the state variables, which are the age and the condition, have reached some threshold, whichever occurs first.

This rule depends on the ratio between the failure replacement cost and the preventive replacement cost. According to the theory of renewal reward processes, the expected cost per unit time can be expressed as:

$$\frac{\text{preventive replacement cost}*\text{ preventive replacement probability}+\text{failure replacement cost}*\text{ failure replacement probability}}{\text{expected replacement time}} \quad (1)$$

It has been shown that the threshold d^* is the optimal cost at which expected cost per unit time is minimal. The optimal threshold d^* can be found by using the fixed-point iteration procedure. Similarly, we can represent the availability function as:

$$\frac{uptime}{uptime+downtime} \quad (2)$$

The optimal availability is achieved by finding the expected time between replacements in the optimal strategy, which depends on the time required to perform the preventive replacement, and the time required to perform failure replacement. The important question in tool replacement strategy is “Should we keep running or should we replace the tool now?” To answer this question, we need to monitor the tool wear and age at discrete time intervals and compare them with a decision rule.

We note that in both strategies of optimal cost and optimal availability, it is expected that some failure will still occur. This means that even if a preventive maintenance is adopted, some failures will still occur because of the randomness nature of the degradation and the failure in most machines.

Experiment description

In defining the decision rule in a practical setting, we set up an experiment using titanium metal matrix composites. A cylindrical bar of Ti-6Al-4V alloy matrix reinforced with 10-12% volume fraction of TiC ceramic particles is used. TiSiN-TiAlN nano-laminate PVD coated grades (Seco TH1000 coated carbide grades) are utilized as cutting tools. We used a 6-axis Boehringer NG 200, CNC turning center in order to conduct experiments. Based on the recommendation of the tool supplier, the experiments have been conducted under the following constant cutting conditions: Cutting speed equals 60 m/min, feed rate equals 0.15 mm/rev, and depth of cut equals 0.2 mm.

Sequential inspections and turning tests are conducted for each tool in order to measure the tool flank wear. The wear is measured after each inspection by using an Olympus SZ-X12 microscope. The procedure continues until the tool flank wear threshold is equal to 0.2 mm. The procedure is replicated for six tools. The collected data is shown in Figure (4-7), where each curve represents one of the six tools, and each point represents an inspection reading. The cutting tool failure is defined by the tool flank wear length reaching a predefined threshold equal to 0.2 mm

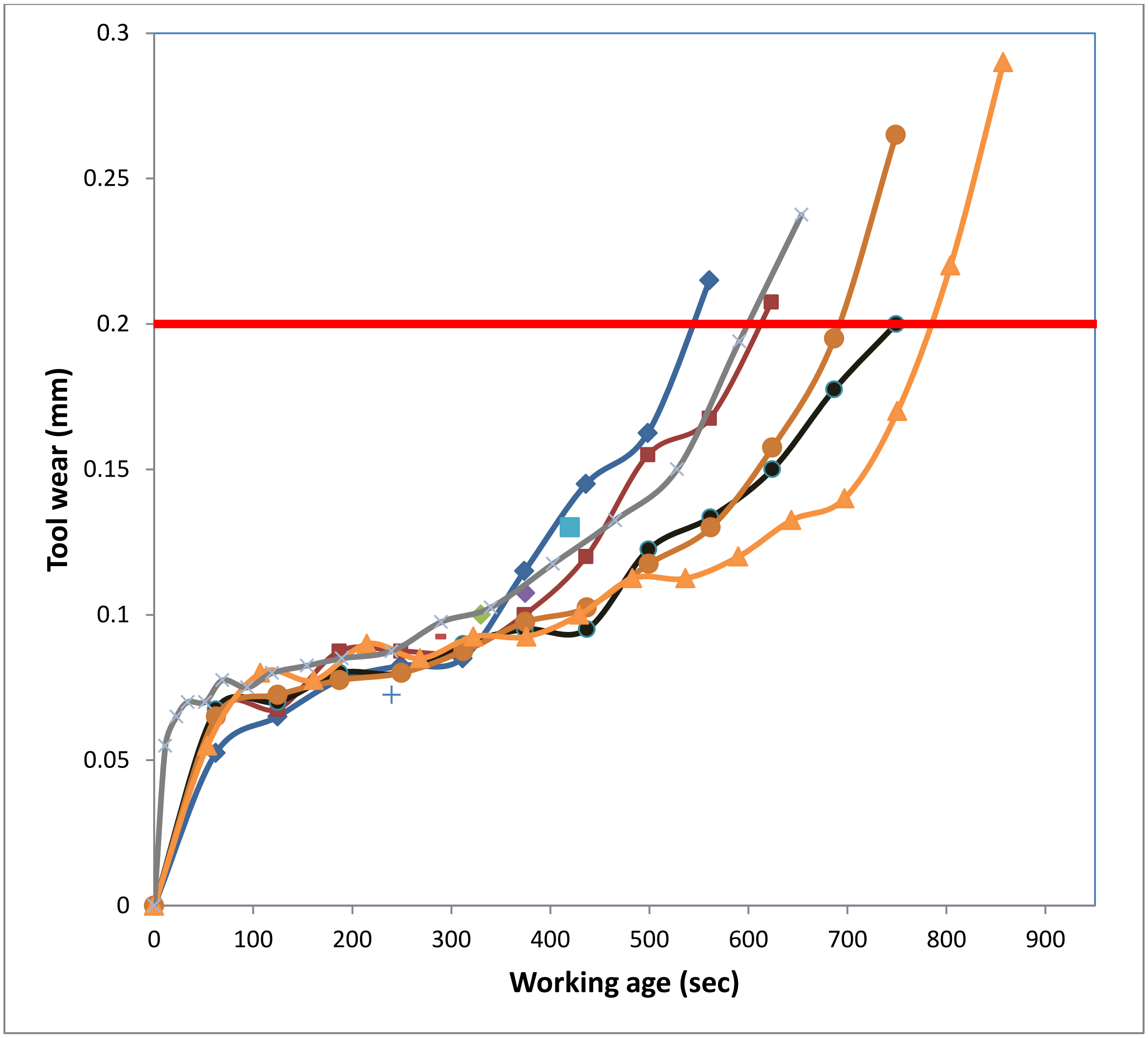


Figure 4-7: Tool wear measurements for 6 tools

In order to calculate the time to failure, the wear evolution between two measurements around 0.2 mm is assumed to be linear. The time to failure is found by interpolating at tool wear equals 0.2 mm. For example, from Table (4.6), and by interpolating between the fifteenth and the sixteenth inspections, the time to failure is found to be 782.73 sec. This interpolation is repeated for the six tools. The times to failure for the six tools are equal to 542.97, 599.94, 611.61, 691.14, 749.17, and 782.73 sec respectively.

Table 4.6: The experimental results showing the wear of tool number 6

Inspection No	Time (sec)	wear(mm)
1	0	0
2	53.7	0.055
3	107.4	0.08
4	161.1	0.0775
5	214.8	0.09
6	268.6	0.085
7	322.3	0.0925
8	376	0.0925
9	429.5	0.1
10	483	0.1125
11	536.5	0.1125
12	590	0.12
13	643	0.1325
14	697	0.14
15	750	0.17
16	804.1	0.22

Development of the model and results

By using the software EXAKT, PHM is constructed by statistically analyzing the wear data, along with the corresponding age of the tools that were removed due to failure. EXAKT is originally designed to be integrated into a plant's maintenance information system to optimize its condition based maintenance (CBM) activities. In 1997, A. K. S. Jardine and V. Makis at the University of Toronto developed the first version of EXAKT, and rapidly earning attention as a CBM-optimizing software. The PHM parameters are estimated, and the resulting hazard function is given as the following equation:

$$\frac{3.713}{86330}\left(\frac{time}{86330}\right)^{2.713} e^{109.1*wear}$$

Hazard function is also called hazard rate or failure rate. It is defined as the instantaneous potential per unit time for the failure to occur, given that the tool has survived up to certain time. This means that Hazard function is a probability of the failing at any given instance. The hazard is a rate rather than a probability and it takes value between 0 and infinity, and depends on whether time is measured in seconds, minutes, hours. The important question in forming our decision policy is “which hazard level should we intervene?”. We wish to choose a hazard level intervention point that results in low cost or high availability.

EXAKT offers Kolmogorov-Smirnov test to evaluate the model fit. The test shows that the PHM offers a good modeling for the data. After determining the PHM, the optimal replacement strategy-cost analysis is performed. The expected time between replacements in the optimal strategy is calculated with a cost ratio of 2:1, preventive replacement cost is estimated to be $100, and the failure replacement cost is $200, thus the difference between the failure and preventive replacement costs is equal to $100. Table (4.7) summarizes the results of the cost optimization’s analysis. It compares the optimal cost per unit time, 0.154 $/sec, which is the quotient of division of equation (1). The numerator of equation (1) equals ($100*0.892+$ 200*0.108); the denominator of equation (1) equals 719 sec, which is the expected time between replacements, in the optimal strategy, with those (0.255 $/sec and 784 sec) of the "run to-failure" strategy, that is without any preventive maintenance. The optimal strategy quantifies the expected preventive and failure costs (0.124 $/sec and 0.03/sec respectively) and the percentage of incidences (89.2% of preventive actions and 10.8% of failure replacement action) achieved when this optimal strategy is used. The expected preventive replacement cost (0.124$/sec) is calculated by dividing the product of preventive replacement cost and preventive replacement probability (100*0.892) by the expected time between replacements (719). Similarly, the failure replacement cost (0.03

$/sec) in column 4, Table (4.7), is calculated. Finally, the table shows that the optimal strategy proposes more interventions; on the average every 719 sec for the optimal strategy versus 784 sec for the strategy of ‘run- to- failure, in order to achieve a net per unit time saving (of $0.101 or 40%). The optimal intervene hazard is found to be 0.154076 /sec in cost analysis. We note again that in any preventive maintenance strategy, there will be some unplanned replacements due to failure. In an optimal strategy, these incidences of failure replacement and the incidence of preventive replacements represent the strategy that will lead to the minimal cost per unit time.

Table 4.7: Summary of cost analysis

	Cost [$/sec]	Preventive Replacement Cost [$/sec]	Failure Replacement Cost [$/sec]	Preventive Replacement [%]	Failure Replacement [%]	Expected time between Replacements
Optimal Strategy	0.154	0.124 (80.5%)	0.03 (19.5%)	89.2	10.8	719
Replacement Only At Failure	0.255	0 (0%)	0.255185 (100%)	0	100	784
Saving	0.101 (40%)	-0.124033	0.225142	-89.2	89.2	-65

Similarly, the availability analysis relies on replacement times. It’s assumed that replacement costs are negligible. The expected time between replacements is calculated when the time required to preventive replacement equals 160 sec, and the time required to failure replacement equals 540 sec. From the results shown in Table (4.8), it is found that the optimal percentage availability (up time=664.68 sec) is equal to 78.75% of the total uptime and downtime (843.994 sec) as in equation (2) , and the percentage of incidences (94.9% of preventive actions and 5.1% of failure replacement actions) achieved when the optimal strategy is used, while the percentage availability (uptime=784 sec) is 59.21% of the total uptime and downtime (1323.75sec), respectively, in the "run to-failure" strategy. The down time is the sum of two parts; the first part is preventive replacement time multiply by preventive replacement probability, and the second term is failure replacement time multiply by failure replacement probability. For example, in optimal strategy, the down time equals 179.318 which calculated by (160*0.949+540*0.051). The percentage of preventive down time (17.99%) in column 3, Table (4.8), is calculated when the product of preventive replacement time and preventive replacement probability (160*0.949) is dividing by the total uptime and downtime (843.994 sec). The percentage of preventive down time can also be considered as 84.69% when the product of preventive replacement time and preventive replacement probability (160*0.949) is dividing by only the down time (179.318). Similarly, the failure downtime percentages in column 4, Table (4.8), are calculated. These results represent an availability savings of 19.54%. Practically speaking, we "buy" high availability by paying for it with more frequent interventions. The optimal intervene hazard is found to be 0.269 /sec in availability analysis.

Table 4.8: Summary of availabilty analysis

	Availability [%]	Preventive Downtime [%]	Failure Downtime [%]	Preventive Replacement [%]	Failure Replacement [%]	Expected time between Replacements [s]
Optimal Strategy	78.75 (664.68*)	17.99 (84.69%)	3.25 (15.31%)	94.9	5.1	843.994 (179.318**)
Replacement Only At Failure	59.21 (784*)	0 (0%)	40.79 (100%)	0	100	1323.75 (540**)
Saving	(19.54%)	-17.99	37.54	-94.9	94.9	-479.752

* expected uptime, **expected downtime

In practice, the preventive replacement cost is equal to $100, the failure replacement cost is equal to $200, and the planned inspection interval equals 60 sec. The PHM model parameters are considered collectively in order to build the “warning level” function as shown in Figure (4-8). Once the decision model is built, we can make a decision that will optimize the long-run maintenance cost for the tool, or the long run availability of the tool. By defining the tool working age and the composite covariate, the optimal decision is to determine whether the tool should be replaced immediately (the red area in Figure (4-8)), or kept operating until the next inspection time (the green area), or kept operating but expected to be replaced before the next inspection time (the yellow area).

Moreover, the model was examined by using historical data to see what the decision model would have recommended for failed tool. The data in Table (4.6) for tool 6 is shown in Figure (4-8). The decision chart gives us the alert of "intervene immediately" at working age 750.63 sec (inspection number 15 in Table (4.6)) because the composite covariates is equal to $109.1 * \text{wear} = 109.1 * 0.17 = 18.55$ when cutting time equals 750.63 sec. This point crosses the "warning level" function. Obviously, in this case, the model was capable of predicting the best action to take. The optimal replacement decision gives a 'warning alert' before the tool's failure. We recapitulate the optimal decision strategy in following words "the optimal strategy suggests replacement at time t for which $109.1 * \text{wear} \geq 34.41 - 2.713 \ln \text{t}$ ".

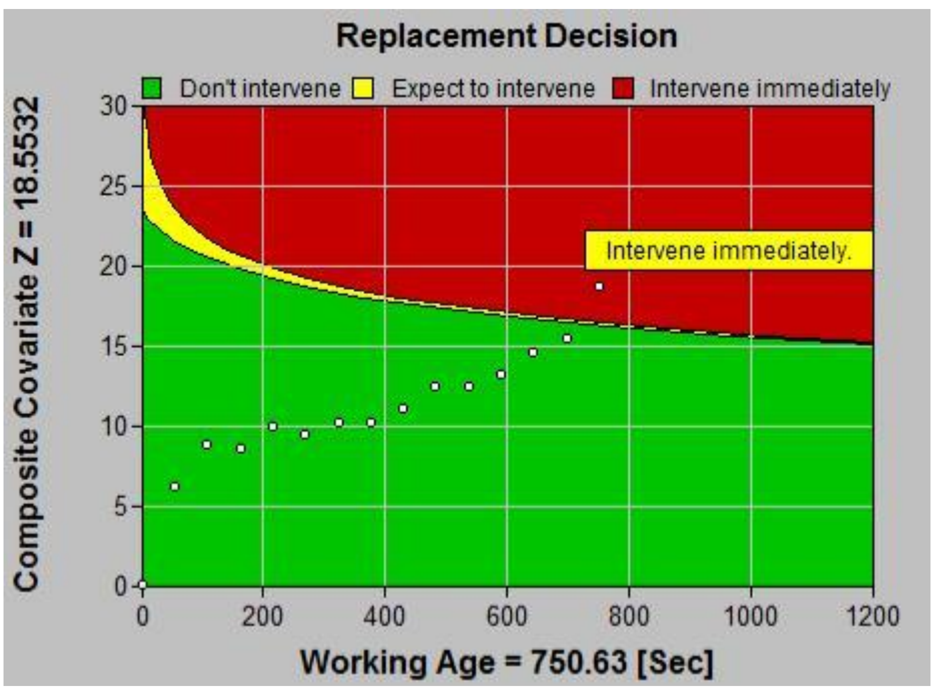


Figure 4-8: Condition-based replacement strategy-optimal decision.

In summary, we collected experimental data during turning titanium metal matrix composites (TiMMCs) to construct the PHM model which was then used to find optimal tool replacement time. The PHM offered a good modelling for the times to failure and tool wear degradation. The

PHM models' parameters and economic objectives were considered to build the optimal decision chart. The study concluded that the optimal replacement times either lead to a cost reduction of 40 percent in case of cost analysis or lead to an increase of 79 percent in the case of availability analysis. Since sensor and information technologies are both expanding rapidly and continuously, it is expected that these analyses will reduce the cost of the machining process. Results from this work are expected to have an impact in the future of optimization of machining processes, especially in biomedical, aerospace and aviation industries.

In future work, the tool wear will be monitored either directly by using a coupled device (CCD) camera which will follow the evolution of tool wear on-line, or indirectly by predicting the wear by monitoring the machining forces, and then by using a machine learning technique, the data will be used for decision making.

CHAPTER 5 ARTICLE 3: OPTIMAL REPLACEMENT TIMES FOR MACHINING TOOL DURING TURNING TIMMCS UNDER VARIABLE MACHINING CONDITIONS

Y. Shaban, M. Aramesh, S. Yacout, M. Balazinski, H. Attia, H. Kishawy

Accepted in:

Journal of Engineering Manufacture

5.1 Abstract

Little practical results are known about the cutting tool optimal replacement time, specifically for machining of composite materials. Due to the fact that tool failure represents about 20% of machine down-time, and due to the high cost of machining, in particular when the work piece's material is very expensive, optimization of tool replacement time is thus fundamental. Finding the optimal replacement time has also positive impact on product quality in terms of dimensions, and surface finish. In this paper, two new contributions to research on tool replacement are introduced. First, tool replacement mathematical models are proposed. These models are used in order to find the optimal time to tool replacement when the tool is used under variable machining conditions, namely the cutting speed, and the feed rate. Proportional Hazards Models (PHM) are used to find an optimal replacement function. Second, this model is obtained during turning titanium metal matrix composites (TiMMCs). These composites are a new generation of materials which have proven to be viable in various industrial fields such as biomedical and aerospace, and they are very expensive. Experimental data are obtained and used in order to develop and to validate the PHM models, which are then used to find the optimal replacement conditions.

Keywords

Optimal tool replacement, metal matrix composites, cost optimization, availability optimization.

5.2 Introduction

Ti-MMCs inherit outstanding characteristics such as low weight, high mechanical and physical properties, high stiffness and strength. Although very expensive, MMCs are a new generation of materials which have proven to be viable in various fields such as biomedical and aerospace industrial. Finding the optimal tool replacement time in machining Ti-MMCs is important in order to decrease the scrapped products and thus the cost of machining, and/or to increase the tool life, and thus to increase the availability of the cutting tool. Replacing the tool only at failure may leave undesired effects on the product's quality characteristics, namely the dimensions and the surface finish. This may lead to scrapping the product. The poor tool condition may cause the waste of subsequent production resources and the loss of customer's goodwill (Hui and Leung

1994). In general, the determination of the optimal replacement time is considered an important economic factor in machining (Klim et al. 1996).

The cutting tool cost represents around 25 percent of the total machining cost (Sakharov et al. 1990; Gray et al. 1993). The cutting tool failure represents about 20% of machine down-time (Liang et al. 2004), replacing cutting tool earlier or later than necessary will cause either loss of valuable resources or products may be scrapped (Tail et al. 2010). Moreover, the tool replacement policy is one of the important aspects of tool management (Jeang 1998). For these reason, finding the time at which the tool should be replaced is fundamental. Much research tried to improve tool life in several ways. For example, Klim et al (Klim et al. 1996) proposed a method to improve cutting tool life in machining using the effect of feed variation on tool wear and tool life. By changing feed rate, the reliability function is changed, and thus the tool life is changed. The Weibull distribution was used to fit the data. The experiment was conducted under constant cutting speed. Balazinski and Mpako (M Balazinski and Mpako 2000) proposed an improvement of tool life through using two discrete feed rates. The method depends on varying the feed rate throughout the cutting process. By varying the feed, the tool-chip contact area increases, the tool wear rate decreases and consequently leads to improvement of the cutting tool life. The experiment was conducted under constant cutting speed. Lin and Shyu (Lin and Shyu 2000) concluded that using variable feed machining, and constant cutting speed, when drilling stainless steel is a significant method for improving the cutting tool life.

Other researches tried to find the optimal replacement strategy by using PHM for modelling tool life, then using another technique to find optimal strategy. For example, Mazzuchi and Soyer (Mazzuchi and Soyer 1989) used a PHM to assess machine tool reliability. Fully Bayesian analysis is used to find optimal machining conditions. Liu and Makis (H. Liu and Makis 1996) derived a formula to calculate the cutting tool reliability under variable cutting conditions. They used PHM while considering the machining conditions as covariates. In(P. H. Liu et al. 2001), the work was extended by developed algorithm based on stochastic dynamic programming for finding the optimal tool replacement times in a flexible manufacturing system. Ding and He (Ding and He 2011) used a PHM by considering vibration signals as a time–dependent covariate. The author suggests that vibration signals are good indicators to tool wear. Reliability analysis based on feature extraction from tool vibration signals is introduced. They found remarkable relationship between the tool condition monitoring information and the life distribution of tool wear by using PHM. Other research used classical Weibull distribution to fit tool life distribution. For example, In (Vagnorius et al. 2010), the Weibull distribution is used to fit tool life distribution. The optimal replacement time for metal cutting is determined from a total time on test (TTT) plot.

Some researchers tried to improve the cutting tool life by changing feed rates while the cutting speed is constant (Klim et al. 1996; M Balazinski and Mpako 2000; Lin and Shyu 2000), others consider the PHM as good model for tool life representation(Mazzuchi and Soyer 1989; V Makis 1995; Tail et al. 2010). In most of these models, it was assumed that the machining conditions have significant effect over the entire tool life but finding tool replacement models is still unavailable. The objective of this paper is to find tool replacement optimization models which can be used in order to minimize the cost or maximize the availability during turning titanium metal matrix composites (TiMMCs) under variable conditions. The PHM is used to model in order to find these models. The Cutting speed (v) and the feed rate (f) are treated as the models' covariates. In section 2, a brief description of the PHM of a tool operating in variable conditions is introduced. In section 3, the optimal replacement policy for minimizing the cost and maximizing the availability is described. In section 4, the experimental procedure which was carried out in order to collect data that is used for constructing the model is presented. The model developed and the final results are presented in section 5. Practical use and sensitivity analysis are given in section 6. Concluding remarks are given in section 7.

5.3 Model description of a tool operating in varying conditions

In 1907, Taylor (Taylor 1907) developed the classical relationship between tool-life (T) and cutting speed (v).The Taylor tool life equation is $v\,T^n = K$, where K and n are experimental constants which depend on the machining conditions, the material of cutting tool and the part. The Taylor's equation shows that the tool-life is inversely proportional to cutting speed. Taylor's extended equation including machining conditions namely, the cutting speed v and the feed f is given in(Mazzuchi and Soyer 1989).This equation has the following form :

$$T = C/(\,v^x\, f^y\,) \tag{1}$$

Where C, x, y are positive constants. Taylor's extended equation considers only the machining parameters but fails to consider the aging and the progressive wear of the tool's effect (Mazzuchi and Soyer 1989). In order to take into consideration the tool's age, the tool life, T is considered a random variable. Due to the flexibility of the Weibull distribution, it is extensively used in modelling the tool life. The Weibull failure rate for a tool in constant operating conditions, that is the speed and feed, is given as follows:

$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \tag{2}$$

Where β is the shape parameter, η is scale parameter, In PHM , the failure rate of the cutting tool is not only dependent the age of the tool, but is also affected by covariates which describe the machining conditions(Mazzuchi and Soyer 1989) . Based on (2), The PHM consists of the failure rate as the product of a baseline failure rate $h_0(t)$, which is dependent only on the age of the tool, and on an exponential expression which is the linear sum of $\gamma_i Z_i$, Z represents the covariates of the machining conditions. The failure hazard rate at time(t) is expressed as in equation (3):

$$h(t, Z; \beta, \eta, \gamma) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\{\textstyle\sum_1^m \gamma_i Z_i\} \tag{3}$$

Using the Weibull model as a baseline function in modelling the tool failure was considered in (V Makis 1995; Tail et al. 2010; Mazzuchi and Soyer 1989). This model is sometimes called the Weibull parametric regression model. The covariates are the cutting speed (v) and the feed rate (f). The model is given in equation (4), where m=2,

$$h(t, Z; \beta, \eta, \gamma_1, \gamma_2) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma_1 v + \gamma_2 f} \tag{4}$$

In this paper, we consider two sates; the normal and the failure states. This latter is defined by the tool wear reaching a predefined level $VB_{Bmax} = 0.2\ mm$ .The survival function can thus be given as in equation (5),

$$R(t; Z) = P(T > t|Z) = \exp\{-H(t, Z)\} = \exp\left\{-\int_0^t h(t, Z)dt\right\} = \exp\{-\left(\frac{t}{\eta}\right)^{\beta} e^{\gamma_1 v + \gamma_2 f}\} \tag{5}$$

Where $H(t, Z)$ is the cumulative hazard function. The survival function $R(t; Z)$ and its derivative $\dot{R}(t; Z) = h(t, Z)R(t; Z)$ are used to estimate the parameters $(\beta, \eta, \gamma_1, \gamma_2)$ by using maximum likelihood (ML) function(Banjevic et al. 2001).

5.4 Optimal replacement Policy

The classical age replacement strategy recommends replacement the cutting tool at failure, that is when the tool wear threshold is reached, or when it reaches a certain age which minimizes the cost per unit time. In the classical strategy, the effect of the covariates are not taken into account. In this paper, the effect of the cutting speed (v) and the feed rate (f) are taken into consideration. The failure hazard rate of the cutting tool is a non-decreasing monotonic function, so the control-limit is used to find the minimum expected cost per unit time(Aven and Bergman 1986; Viliam Makis and Jardine 1992). The control-limit is a control-limit value $(d > 0)$. The optimal stopping rule is given in equation (6). The stopping rule is often used in condition based maintenance CBM as an alarm when uncontrollable covariates reach predefined states. In this paper, it is used As follows:

$$T_d = inf\{t \geq 0 \colon Kh(t,Z) \geq d\} \tag{6}$$

Where T_d is the preventive replacement time, K is the difference between the failure replacement cost $C + K$ and the preventive replacement cost C. According to the theory of renewal reward processes, the expected cost per unit time can be expressed as

$$\phi(T_d) = \frac{C\,P(T_d<T)+(C+K)\,P(T_d\geq T)}{W(d)} = \frac{C+K\,P(T_d\geq T)}{W(d)} \tag{7}$$

$d^* = \phi(T_d^*)$ is the optimal cost at which the $\phi(T_d)$ *is minimum* and T_d^* is the optimal time to replace. $P(T_d \geq T)$ is the probability of failure replacement, $P(T_d < T)$ is the probability of preventive replacement, and $W(d) = E(\min\{T_d, T\})$ is the expected replacement time. Optimal level d^* can be found by using the fixed-point iteration procedure (Banjevic et al. 2001; Viliam Makis and Jardine 1992) or by using Semi-Markovian Covariate Process(Bergman 1978).

Similarly, we represent the availability function as in equation (8).

$$A(T_d) = \frac{uptime}{uptime+downtime} = \frac{W(d)}{W(d)+T_p\,P(T_d<T)+(T_p+K)\,P(T_d\geq T)} \tag{8}$$

When $A(T_d)$ is the availability. The optimal availability is achieved at T_d^* the optimal time to replacement, T_p is the time required to perform the preventive replacement, and $T_f = (T_p + K)$ is the time required to perform failure replacement. We note that in equation (8), K is the difference between T_f, T_p, while in equation (6) it is the difference between the failure replacement cost and the preventive replacement cost.

The objective is to find d^*.The replacement function is derived when d^* is obtained and the machining conditions, namely the cutting speed (v), and the feed rate (f) are known. The replacement function is derived from equation (6) as follow:

$$Kh(t, Z) \geq d^* \quad (9)$$

$$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \mathrm{e}^{\gamma_1 v + \gamma_2 f} \geq \frac{d^*}{K} \quad (10)$$

$$\mathrm{e}^{\gamma_1 v + \gamma_2 f} \geq \frac{d^* \, \eta^{\beta} t^{-(\beta-1)}}{K\beta} \quad (11)$$

$$\gamma_1 v + \gamma_2 f \geq \ln\left(\frac{d^* \, \eta^{\beta}}{K\beta}\right) - (\beta - 1)\ln t \quad (12)$$

$$Z^c \geq \mathrm{g(t)} \quad (13)$$

g(t) was defined as a warning function in (Banjevic et al. 2001). The function g(t)= $\ln\left(d^* \, \eta^{\beta} / K\beta\right) - (\beta - 1)\ln t$ can be consider as “replacement” function. By calculating an “overall” covariate value Z^c, the optimal time to replacement T_d^* is obtained.

5.5 Description of the Experiment

Equipment: A 6-axis Boehringer NG 200, CNC turning center is used in order to conduct experiments, as shown in figure (5-1). *Tool material*: TiSiN-TiAlN nano-laminate PVD coated grades (Seco TH1000 coated carbide grades) is used. *Workpiece material*: A cylindrical bar of Ti-6Al-4V alloy matrix reinforced with 10-12% volume fraction of TiC ceramic particles is used. *Experimental details*: The experiments were conducted using full factorial designs with two-factors, two-level ($v = 40, 80\ m/min$ and $f = 0.15, 0.35\ mm/rev$), and using one center point ($v = 60\ m/min$ and $f = 0.25\ mm/rev$). Full factorial designs are the most conservative of all design types because we try all combinations of the factor settings. Table (5.1) shows the design of the experiment in a coded form. Table (5.2) shows all combination of cutting conditions. There are 5 runs which were done randomly. Each run was replicated at least 5 times.

Table 5.1 : The coded design of experiment

Run \ Factor	Cutting Speed	Feed Rate	Depth of Cut
1	-1	-1	1
2	1	-1	1
3	-1	1	1
4	1	1	1
5	0	0	0

Table 5.2: The design of experiment

Run \ Factor	Cutting Speed (m/min)	Feed Rate (mm/ rev)	Depth of Cut (mm)
4	80	0.35	0.2
3	40	0.35	0.2
1	40	0.15	0.2
2	80	0.15	0.2
5	60	0.25	0.2

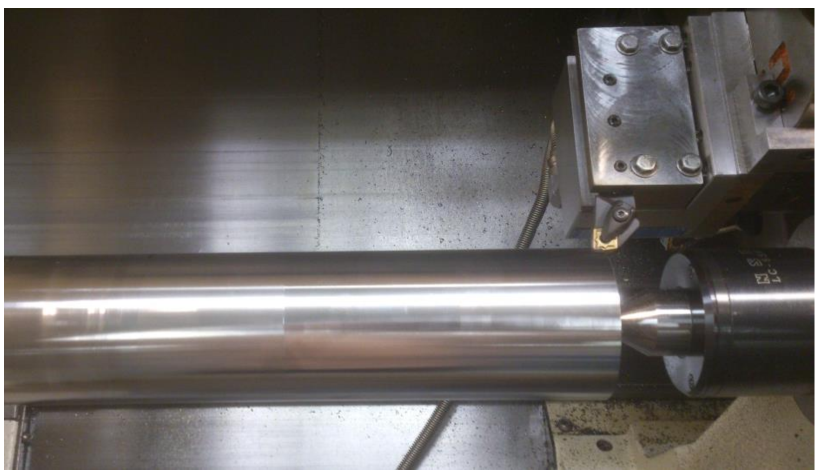


Figure 5-1:The experimental setup

The cutting tool fails when the tool becomes dull and no longer operates within acceptable quality(Gray et al. 1993). The common way of quantifying the tool time to failure is to put a limit on the maximum acceptable flank wear, $\mathrm{VB_{Bmax}}$. For each tool, sequential inspections were conducted in order to measure the wear. The wear is monitored at discrete points of time through inspections. The wear is measured after each inspection by using an Olympus SZ-X12 microscope. The procedure continues until the tool wear threshold ($\mathrm{VB_{Bmax}} = 0.2\ mm$) is reached. The procedure is replicated for 28 tools.

Figure (5-2) shows the wear interpolation procedure in order to calculate the time to failure TTF, the wear evolution between two measurements (VB_i , VB_{i+1}) is assumed to be linear. TTF is calculated when tool wear threshold ($\mathrm{VB_{Bmax}} = 0.2\ mm$) is reached. For example, from Table (5.3), by interpolating between the fourteenth inspection at ($\mathrm{t_i} = 1530\ sec$) and the fifteenth inspection at ($\mathrm{t_{i+1}} = 1650$ sec), and by using equation (14), the time to failure is found to be 1623.3 sec. This interpolation is repeated for 28 tools. The results for the 28 tools are given in Table (5.4).

$$\frac{\varepsilon}{\Delta t} = \frac{0.2 - VB_i}{\Delta VB} \quad , \ TTF = \varepsilon + t_i \qquad (14)$$

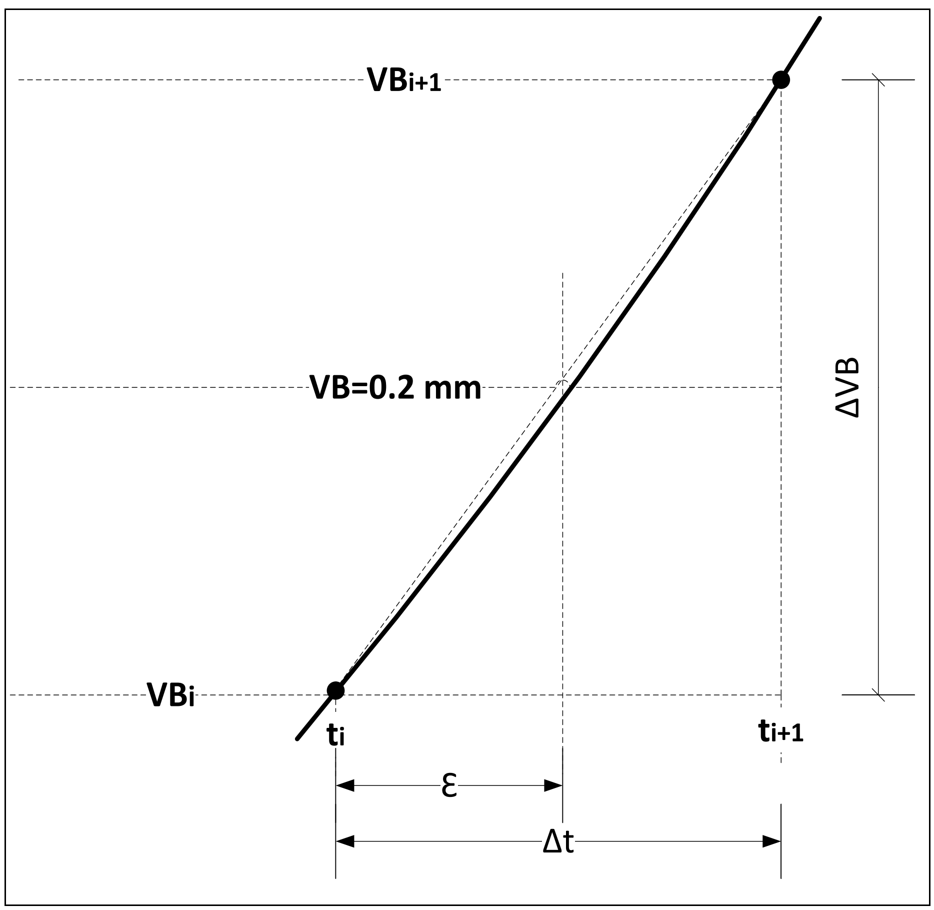


Figure 5-2:Wear interpolating

Table 5.3:The experimental results showing the wear of tool 1-1

Inspection No	Time(sec)	*VB*(mm)
1	0	0
2	120	0.0525
3	240	0.06
4	360	0.065
5	480	0.0725
6	600	0.0875
7	720	0.1075
8	840	0.1125
9	960	0.12
10	1050	0.125
11	1170	0.135
12	1290	0.165
13	1410	0.175
14	1530	0.1825
15	1650	0.205

Table 5.4: Times to failure *TTF* for the 28 tools

Tool ID Run-replication	Time to failure sec	Speed($\boldsymbol{v}$) m/min	Feed($\boldsymbol{f}$) mm/rev	Tool ID Run-replication	Time to failure sec	Speed($\boldsymbol{v}$) m/min	Feed($\boldsymbol{f}$) mm/rev
1-1	1623.3	40	0.15	**3-5**	1230	40	0.35
1-2	2087	40	0.15	**3-6**	1006	40	0.35
1-3	1770	40	0.15	**4-1**	121.4	80	0.35
1-4	1524	40	0.15	**4-2**	87.5	80	0.35
1-5	1560	40	0.15	**4-3**	135	80	0.35
2-1	295	80	0.15	**4-4**	135	80	0.35
2-2	267.5	80	0.15	**4-5**	121.7	80	0.35
2-3	281.2	80	0.15	**4-6**	102.5	80	0.35
2-4	225.3	80	0.15	**5-1**	233.5	60	0.25
2-5	252.7	80	0.15	**5-2**	192	60	0.25
3-1	1240	40	0.35	**5-3**	265	60	0.25
3-2	1002	40	0.35	**5-4**	190	60	0.25
3-3	1320	40	0.35	**5-5**	160	60	0.25
3-4	1263.3	40	0.35	**5-6**	185	60	0.25

5.6 Development the model and results

The PHM parameters are estimated using Exakt software (Banjevic et al. 2001). The resulting hazard function is given as follows in equation (15):

$$h(t,Z) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma_1 v+\gamma_2 f} = \frac{3.71}{23760}\left(\frac{t}{23760}\right)^{2.71} e^{0.195v+10.86f} \tag{15}$$

The covariate parameters γ_1= 0.195 and γ_2= 10.86 are the multipliers for cutting speed (v) and the feed rate (f) respectively in the hazard function. A small value for γ_1 parameter does not mean that cutting speed (v) has a small effect on the hazard function because the covariate parameter is multiplied by the covariate value which can be large(EXAKT help Version 4.20.1 2007). In order to distinguish between statistically significant and non- significance covariates, a formal statistical test is needed. In figure (5-3), statistical Wald test shows in column 5 that the cutting speed is more significant than feed rate.

Parameter	Estimate	Sign. (*)	Standard Error	Wald	DF	p - Value
Scale	2.376e+004	-	6174	-	-	-
Shape	3.71	Y	0.6077	19.88	1	0
v	0.1951	Y	0.03356	33.78	1	0
f	10.86	Y	2.735	15.76	1	0

Figure 5-3:Summary of estimated parameters (based on ML method).

In order to know how the cutting speed and the feed rate affect the hazard rate, a simple normalization procedure is done. Since the cutting speed and the feed are in the range $(40, 80)$ and $(0.15, 0.35)$, respectively,the normalization of the "overall" covariate will be as follow:

$$Z^c = \beta_o + \beta_1 x_1 + \beta_2 x_2 = 14.415 + 3.9\, x_1 + x_2, \qquad (16)$$

$$where x_1 = (v - 60)/20,\; x_2 = (f - 0.25)/0.1, \text{ and}$$

$$x_1, x_2 \in [-1,1]$$

$(\beta_o, \beta_1\, and\; \beta_2)$ are called regression coefficients (Montgomery 2007). In our model, it is obvious that the effect of cutting speed on cutting tool life is approximately four times more than the effect of feed rate.

In order to validate the model, Kolmogorov-Smirnov test (K-S test) and logarithmic reliability function analysis are done. (K-S test) evaluates the model fit. The test checks the null hypothesis that the $\mathrm{H(t, Z)}$ in equation (5) is distributed exponentially(P. H. Liu et al. 2001). The summary of goodness of fit test is automatically produced in EXAKT as in table (5.5).The test shows that the PHM offers a good modeling for the data.

Table 5.5: Summary of goodness of fit test results

Test	Observed value	P-value	PHM Fits Data
Kolmogorov-Smirnov	**0.2266**	**0.0965714**	**Not rejected**

Figure (5-4) shows the analysis of the logarithmic reliability function (log minus log plot) (Kalbfleisch and Prentice 2011). From equation (5), the linear equation for each run will be as follow:

$$Ln\left[-Ln\big(R(t;Z)\big)\right] = \beta\, Ln(t) - \beta\, Ln(\eta) + \gamma_1 v + \gamma_2 f \qquad (17)$$

The logarithmic reliability function in equation (17) is linear in $Ln(t)$, and for each run, corresponding functions are parallel (Tail et al. 2010). It is concluded, now, that the PHM-model's assumption is satisfied and presented the reliability functions of the cutting tool in the range of the cutting speed and the feed rate

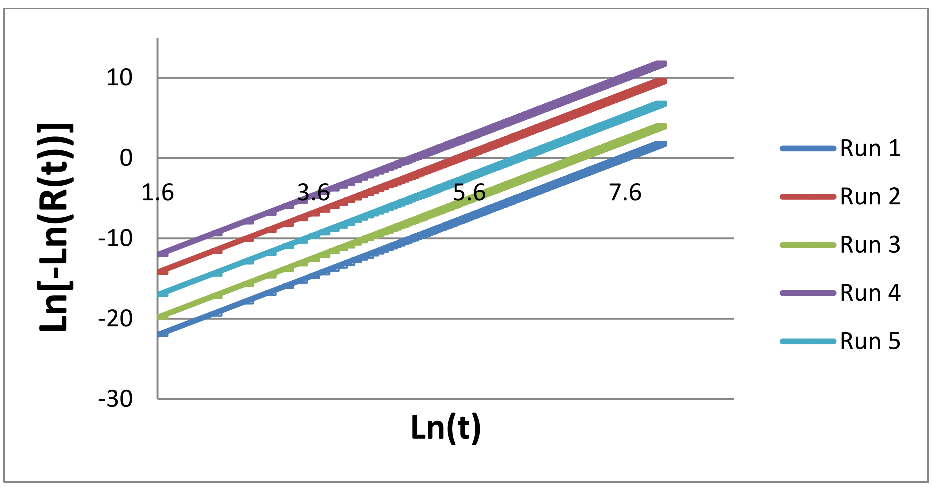


Figure 5-4: logarithmic reliability function plot for each run

Based on equation (15), the failure rates are plotted for each run in figure (5-5). The effect of machining conditions on the failure risk is clear when we compare between different runs. For example, by comparing between run 1 and run 2 which have the same feeds rates and but different speeds, and also by comparing between run 2 and run 4 which have the same speeds and different feed rates, obviously, the effect of cutting speed is much higher than the effect of feed rate.

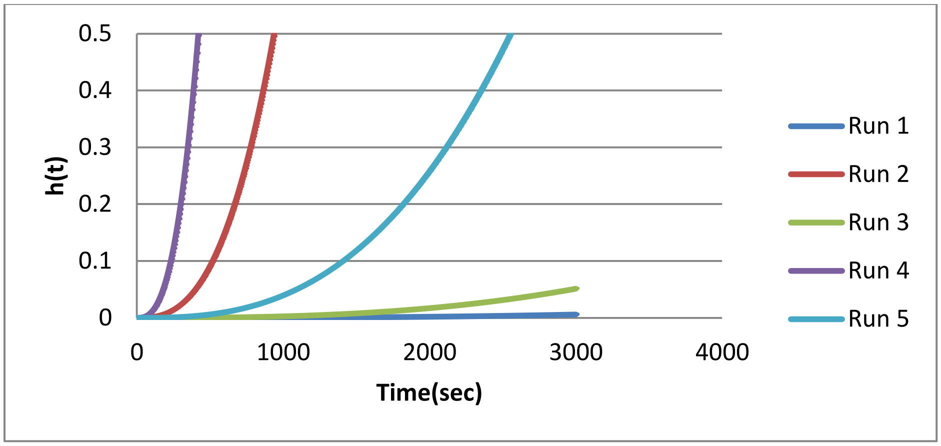


Figure 5-5: Hazard rate curves for each run

After determining The PHM, the optimal replacement policy-cost analysis is performed. The optimal replacement function is calculated with a cost ratio r = 2 (preventive replacement cost is estimated to be \$100, and the failure replacement cost is \$200, thus K is equal to \$100). As shown in figure (5-6), the optimal time to replacement T_d^* can be calculated. The function g(t)= $\ln\left(d^* \eta^\beta / K\beta\right) - (\beta - 1)\ln t = 29.429 - 2.71\, ln\, t$ is the replacement function, applied to an “overall” covariate value $Z^c = 0.195\, v + 10.86\, f$.

Figure 5-6: Optimal replacement function-cost analysis

Similarly, we find the optimal replacement function that maximizes the availability. The optimal time to replacement T_d^* is then calculated. The time required to perform preventive replacement, $T_p = 160$ sec, and the time required to perform failure replacement,$T_f = 540$ sec). As shown in figure (5-7), the function g(t)= $\ln\left(d^* \eta^{\beta}/K\beta\right) - (\beta - 1)\ln t = 31.22 - 2.71\ln t$ is the replacement function, applied to an "overall" covariate value $Z^c = 0.195\, v + 10.86\, f$.

Figure 5-7: Optimal replacement function-availability analysis

In practice, finding optimal replacement policy is generalized. Figure (5-8) shows the sequence of finding the optimal replacement T_d^* in both cases of cost analysis or availability analysis. For example, in cost analysis, the procedure is as follows:

1. Extract the event (tool failure) by sequential inspections for any machining process.
2. Collect the experimental data in order to build the model by estimating the parameters of the PHM model.
3. Check the goodness of fit using, for example, Kolmogorov-Smirnov test.
4. Find $d^* = \phi(T_d^*)$ $which\ is\ the\ optimal\ cost\ where$ $\phi(T_d)\ is\ minimum$, and then find the replacement function, g(t)= $\ln\left(d^* \eta^{\beta}/K\beta\right) - (\beta - 1)\ln t$ for a known costs $C\ and\ C + K$.
5. Calculate T_d^* for current machining conditions (v and f) by defining the composite covariate $Z^c = 0.195\, v + 10.86\, f$ and using the replacement function.

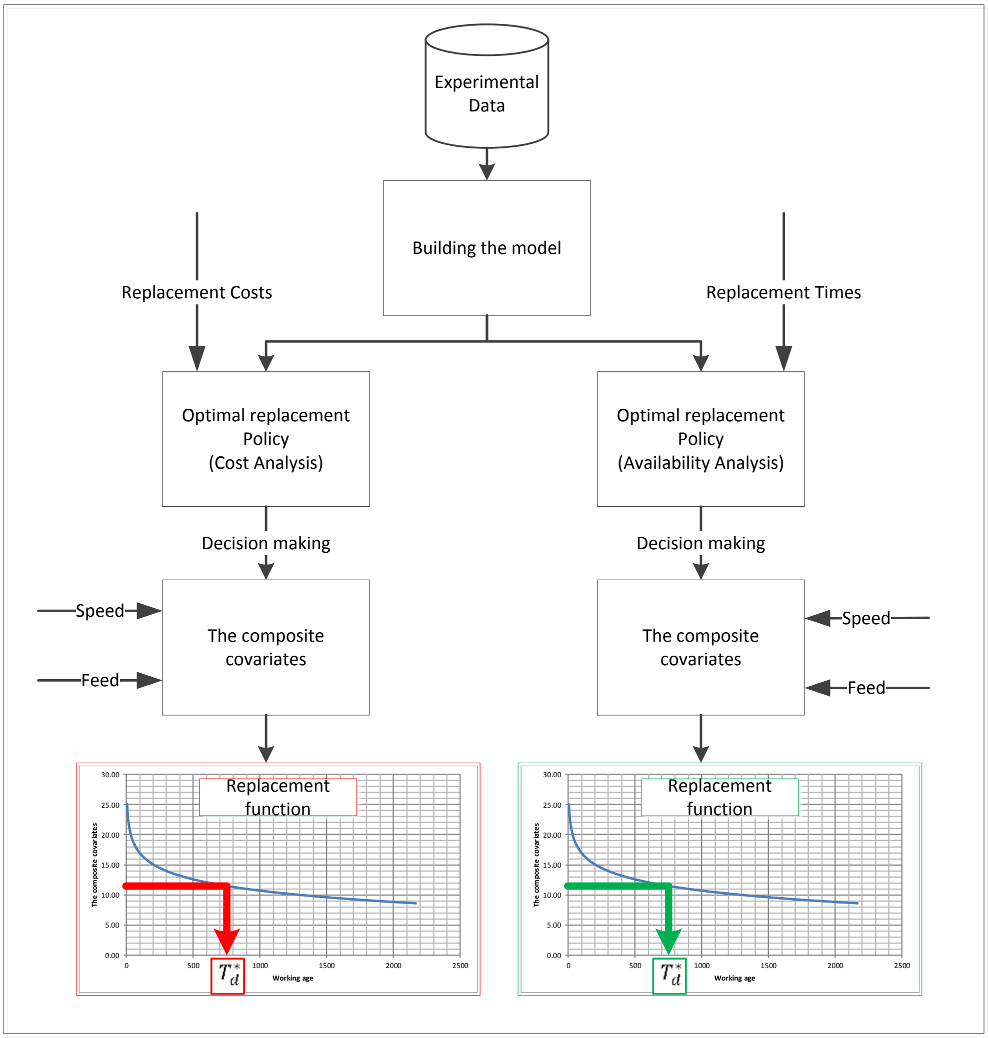


Figure 5-8: Finding the optimal replacement time in cost and availability analysis

5.7 Practical use and sensitivity analysis

The replacement function is used for a single cutting tool in multitasked machining process under variable machining conditions. For example, the user may use the tool for machining a part with machining conditions ($v = 50\ m/min$ and $f = 0.20\ mm/rev$) for 200 sec, then he/she may want to use the same tool for a second machining process with machining conditions ($v = 40\ m/min$ and $f = 0.15\ mm/rev$). The question is "Can he/she use this tool for the second machining process and for how long he/she can use this tool before replacing it with a new one in order to get the cost optimality". Figure (5-9) answers this question. The first machining process starts at point 1 while $Z^c = 11.92$ and continues horizontally until point 2 ($T_d = 200$ sec). Since Point 2 is below the replacement function curve, the user can used the tool for the second machining process which will start at point 3 while $Z^c = 9.43$ and can go horizontally until it touch the replacement function curve (point 4), which gives the optimal time to replacement $(T_d^* = 1610 \text{ sec})$. The optimal remaining time for the second machining process is $(T_d^* - T_d = 1410 \text{ sec})$. Obviously, that example shows how user can follow the status of the cutting tool by knowing its cutting speed (v), feed (f), and working age.

Figure 5-9:Optimal replacement example-cost analysis

Sensitivity analysis is performed on the cost ratio (r). Figure (5-10) shows the cost ratio sensitivity when r = 2 to r = 5. Obviously, the optimal time to replacement is decreasing when the cost ratio (r) is increasing. This is very logical because as the difference between the failure replacement cost and the preventive replacement cost gets higher, the more frequent preventive replacement should be done, thus the new optimal time to replacement will be less than the original one, in order to minimize the cost per unit time.

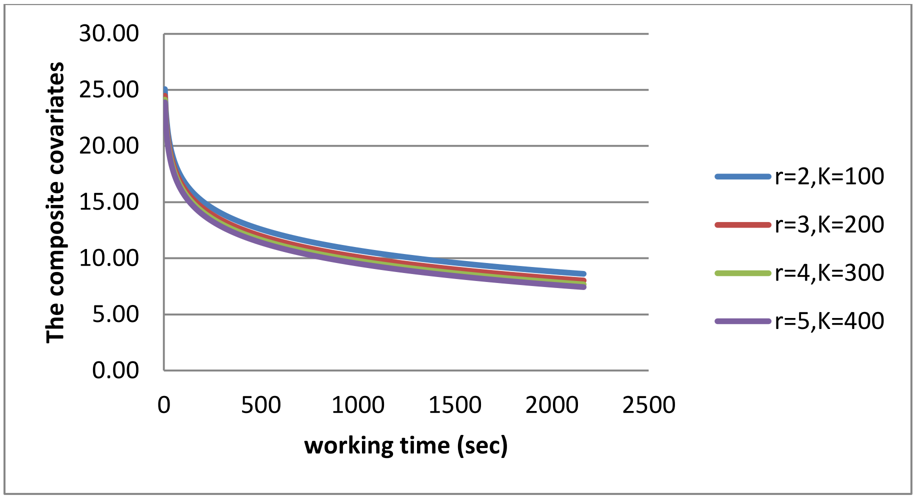


Figure 5-10: The cost ratio sensitivity

5.8 Conclusion

In this paper, we have introduced two new contribution to the research on tool replacement based, which are two optimality models for cost minimization and availability maximization, and we applied it to a new generation of composites, namely the TiMMCs. Experimentally, data were collected during turning titanium metal matrix composites (TiMMCs) under variable machining conditions. The collected data were used to construct the PHM model. The PHM offered a statistically good model for the problem. An optimal replacement function was obtained and built into a simple chart. While changing the machining conditions, we showed how the user can find the optimal time to replacement that optimizes either the machining cost or the availability per unit time.

CHAPTER 6 ARTICLE 4: TOOL WEAR MONITORING AND ALARM SYSTEM BASED ON PATTERN RECOGNITION WITH LOGICAL ANALYSIS OF DATA

Yasser Shaban, Soumaya Yacout, Marek Balazinski,

Submitted to:
ASME Journal of Manufacturing Science and Engineering

6.1 Abstract

This paper presents a new tool wear monitoring and alarm system that is based on Logical Analysis of Data (LAD). LAD is a data-driven combinatorial optimization technique for knowledge discovery and pattern recognition. The system is a non-intrusive on-line device that measures the cutting forces and relates them to tool wear through learned patterns. It is developed during turning titanium metal matrix composites (TiMMCs). These are a new generation of materials which have proven to be viable in various industrial fields such as biomedical and aerospace. Since they are quite expensive, our objective is to increase the tool life by giving an alarm at the right moment .The proposed monitoring system is tested by using the experimental results obtained under sequential different machining conditions. External and internal factors that affect the turning process are taken into consideration. The system's alarm limit is validated and is compared to the limit obtained when the statistical Proportional Hazard Model (PHM) is used. The results show that the proposed system that is based on using LAD detects the worn patterns and gives a more accurate alarm for cutting tool replacement.

Keywords

Tools wear monitoring, metal matrix composites, Logical Analysis of Data, pattern recognition.

6.2 Introduction

In the published literature, tool wear in machining processes is analyzed by two approaches: Firstly, theoretical and numerical approach, such as state space methods and finite element method (FEM), and secondly, data-driven approach, such as artificial neural network (ANN) and fuzzy logic (Shi and Gindy 2007). Li (Li 2012) presented an exclusive review of tool wear estimation using theoretical analysis and numerical simulation technologies. Sick (Sick 2002) presented an exclusive review of indirect online tool wear monitoring in turning with ANN as an example of data-driven technique. By indirect, we mean that researchers usually measure covariates (variables) which are indirectly correlated with tool wear such as the cutting forces. These forces are measured on-line during machining process. There are hundreds of researches about tool wear monitoring system. Nevertheless, only a few systems found their way to real industrial application (Jemielniak 1999). The tool wear monitoring systems development is still on-going attempt (Sick 2002). Byrne et al (Byrne et al. 1995) presented a review about utilization of these systems in industry. Another review about commercial tool monitoring systems was done by (Jemielniak 1999).

Due to the availability of sensory signals, data-driven approach has received much attention to build on-line tool wear monitoring systems. Data-driven techniques need training stage to learn how to adjust adaptively to the data without statistical distribution. Once learning stage is accomplished and validated, the system can detect worn pattern correctly. (Damodarasamy and Raman 1993) developed an inexpensive system for classifying tool wear states using pattern recognition. Despite that the accuracy of classification was relatively small; they concluded that pattern recognition can be successfully used to predict the status of cutting tool wear. They combined the feed force, radial force and the root mean square of acoustic emission (AE) signals to predict the tool wear. In (Shi and Gindy 2007), the tool wear predictive model is presented by combination of least squares support vector machines and principal component analysis technique. The platform of PXI and LabVIEW were used to develop the system. (S Purushothaman and Srinivasa 1994) developed a model for classifying a worn-out tool and a fresh tool. They used ANN for building a model. (Kang et al. 2007) developed a method of pattern recognition of tool wear based on discrete hidden Markov models. The results showed that the proposed method is effective. All techniques which used pattern recognition for classifying tool wear states are based on assumptions related to the data structure. In this work, the proposed technique, LAD is not based on any assumptions or statistical techniques. It is used for the first time in tool condition monitoring. In this paper, our objective is to report and discuss the results obtained experimentally.

In the following experiments, the workpiece material is Ti-MMCs which have been well employed in various industrial fields such as biomedical and aerospace. The high strength associated with Ti-MMCs leads to rapid cutting tool wear rate. The poor condition of the tool may leave bad effect over the dimensions and the surface finish of the product causing it to be scrapped. The scrapping of product increases the machining process' cost, especially when the workpiece material is very expensive which is the case of Ti-MMCs. On the other hand, replacing the cutting tool earlier than necessary will cause the loss of valuable resources (Tail et al. 2010). In general, and regardless of the tool replacement costs, the implementation of accurate on-line tool wear monitoring provides a cost-effective solution to the problem of determining the best tool replacement moment.

In this paper, we differentiate between two types of covariates; internal (diagnostic) covariates, which carry direct information about the wear process, and external (environmental or/and machining conditions) covariates, which affect the wear process (W. Wang 2004; Kalbfleisch and Prentice 2011; Banjevic et al. 2001). In machining, external covariates may be controllable, and has predefined determined path such as cutting speed, feed rate, depth of cut, tool geometry, contact angle, tool material, and workpiece material (H. Liu 1997). External covariates may also be uncontrollable such as the ambient temperature and air humidity in the laboratory(W. Wang and Hu 2006). Internal covariates are observed by on-line monitoring of time dependant factors such as cutting forces, cutting temperatures, progressive wear, acoustic emissions and vibration signals. Combination of internal and external covariates was used before in order to develop accurate model. Azouzi and Guillot (Azouzi and Guillot 1997) show that the combination of feed rate, depth of cut, radial force, and feed force in turning process provided accurate model in on-line estimation of surface roughness and dimensional deviations. Their fusion model was built using neural network.

In this paper, we implement tool wear monitoring system based on LAD during turning TiMMCs under variable conditions. The platform of PXI and LabVIEW were used to develop the tool wear alarm system. In section 6.3, a brief description of LAD is introduced. In section 6.4, the experimental procedure which is peformed in order to collect the data that is used for constructing the system is presented. In section 6.5, knowledge extraction and learning from the data is carried out in order to train the system. The Proportional Hazards statistical model (PHM) is presented in section 6.6. In section 6.7, the on-line alarm system that is based on LAD is described, and a comparison with the PHM alarm function is made. Discussion and concluding remarks are given in section 6.8.

6.3 Logical analysis of data (LAD)

LAD is a data-driven combinatorial optimization technique that allows the classification of phenomena based on pattern recognition. LAD is applied in two consecutive stages, training or learning stage, and the testing or the theory formation stage. In learning stage, a part of the data is used to extract special patterns of some phenomena. In testing stage, the remainder of the data is used to test the accuracy of the previously learned knowledge. LAD is based on supervised learning; this means that the data used in the learning stage must be labeled before analysis. That is, each observation belongs to a known class. In this work, we have two classes of cutting tool: worn-out tool, and a fresh tool. Each observation carries the values of the covariates and a label. The covariates are internal and external. The internal covariates are the radial force (f_x), the feed force (f_y), and the cutting force (f_z). The external covariates are the cutting speed (v), and the feed rate (f). After accomplishment of the two phases of learning and testing, worn patterns, which represent the worn-out tool condition, and fresh patterns which represent the normal condition are found by LAD. The worn patterns are used in order to develop tool wear monitoring model. This model is later incorporated in the platform of PXI and LabVIEW in order to monitor the tool wear on-line, and to give an alarm when the tool worn patterns are detected.

In (P.L. Hammer and Bonates 2006) , LAD overview is introduced by the group of researchers at Rutgers University. LAD methodology was compared to other techniques of machine learning (P.L. Hammer and Bonates 2006; Soumaya Yacout 2010). It was concluded that LAD has certain advantages over other techniques. For example, since it is a non-statistical approach; it does not need any prior assumptions regarding the posteriori class probabilities. It also has the advantage of giving the user the ability to track back any results (phenomena or effects) to its possible causes. It is often used as two-class classification technique (Bores et al. 2000). The observations are classified as either positive (fresh, π^+, class 1) or negative (worn-out, π^-, class 2). LAD generates collections of patterns which characterizes each class. These patterns represent interactions between variables (internal and external covariates) in each class, fresh or worn out, separately. The patterns are called worn patterns when they describe the worn-out tool condition, and fresh patterns when they describe normal wear condition After the learning stage, LAD can classify any new observations that are not included in the original dataset (Bores et al. 2000).

LAD has three steps: binarization of data, pattern generation, and theory formation. Data binarization is the process of transformation of data into a Boolean database. The binarization step involves the transformation of the training data to binary data using a binarization technique which is discussed in (Bores et al. 2000). This technique substitutes each numerical variable by at least one binary attribute. For example, binarization of a continuous numerical variable A is done by ranking, in ascending order, all the distinct values of the numerical variable A as follows:

$$u_A^{(1)} < u_A^{(2)} < \cdots < u_A^{(q)} \qquad (q \leq Q) \tag{1}$$

Where q the total number of is distinct values of the variable A and Q is the total number of observations in the training set. The cut-points $\delta_{A,j}$, where j is the number of cut points for each variable, are found between each pair of values that belong to different classes. By using equation (2), the cut-points are calculated as follows:

$$\delta_{A,j} = (u_A^{(k)} + u_A^{(k+1)})/2 \tag{2}$$

Where $u_A^{(k)} \in \pi^+$ and $u_A^{(k+1)} \in \pi^-$ or vice versa. A binary attribute b is then formed from each cut-point. Each cut-point $\delta_{A,j}$ has a corresponding binary attribute $b_{\delta_{A,j}}$ with defined value:

$$b_{\delta_{A,j}} = \begin{cases} 1 & \text{if } u_A \geq \delta_{A,j} \\ 0 & \text{if } u_A < \delta_{A,j} \end{cases} \tag{3}$$

The second step of LAD consists of pattern generation. It is the key building block in LAD knowledge extraction. There are many techniques for pattern generation such as enumeration (Bores et al. 2000), heuristics (Peter L Hammer 1986; P.L. Hammer and Bonates 2006), and linear programming (Ryoo and Jang 2009). Here, We follow the pattern generation technique which has been proposed in (Ryoo and Jang 2009), The authors convert the pattern generation problem to a set covering problem which is solved by mixed integer linear programming (MILP). It should be noted that using linear programming to generate patterns does not mean that LAD uses a mathematical linear model in order to separate between the worn and the fresh conditions. The LAD based knowledge extraction technique is highly non-linear, and can extract knowledge from highly non separable data (M.-A. Mortada et al. 2011). A positive (negative) pattern is defined as a conjunction of some of binary attributes which is true for at least one positive (negative) observation and false for all negative (positive) observations in the training data set. The number of binary attributes used to define the pattern is called the degree of a pattern. For example, pattern p of degree ∂ is a conjunction of ∂ attributes. A pattern covers an observation in the training set if and only if it is true for that particular observation (Bores et al. 2000).

Theory formation or testing stage is the final step in the LAD decision model. A discriminant function, such as the one given in equation (4), is formulated to generate a score ranging between -1 and 1. When the output of a discriminant function is a negative value that means that the tested observation belongs to the negative class, and positive otherwise. Zero value means that LAD cannot classify the observation (M.-A. Mortada et al. 2011).

$$\Delta(O) = \sum_{i=1}^{N^+} \sigma_i^+ Z_i^+(O) - \sum_{i=1}^{N^-} \sigma_i^- Z_i^-(O) \tag{4}$$

Where $N^+(N^-)$ is the number of positive (negative) patterns that are generated, $Z_i^+(O)\left(Z_i^-(O)\right)$ is equal to 1 if pattern (i) covers observation O, and is equal to zero otherwise, $\sigma_i^+(\sigma_i^-)$ is the weight of the positive (negative) pattern $p_i^+(p_i^-)$. The weight represents the proportion of observations that are covered by the pattern. High weight indicates that the pattern covers higher number of observation, thus it is a better indicator of the class to which the observation belongs. The calculated value of $\Delta(O)$ of any new observation gives an indication to whether the observation belongs to fresh or worn-out class. In order to measure the accuracy, the quality of classification (ν) is used.

$$\nu = \frac{a+b}{2} + \frac{c+e}{4} \tag{5}$$

Where the values (a) and (b) represent the proportion of observations, positive and negative, which are correctly classified. The values (c) and (e) represent the proportion of observations, positive and negative, which are unclassified.

6.4 Description of the Experiment

The experiment was conducted in the machining laboratory at École Polytechnique de Montréal. As shown in figure (6-2), A 6-axis Boehringer NG 200, CNC turning center is used in order to conduct experiments. A TiSiN-TiAlN nano-laminate PVD coated grades (Seco TH1000 coated carbide grades) is used. A cylindrical bar of Ti-6Al-4V alloy matrix reinforced with 10-12% volume fraction of TiC ceramic particles is used. Cutting forces are measured using 3-component dynamometer. Forces directions during turning are shown in figure (6-1). The signals are passing through Multichannel charge amplifier and then collect by national instruments acquisition board (PXI 1000B). The experiments were conducted using full factorial design with two-factors, two-level speed ($v = 40, 80\ m/min$) and feed ($f = 0.15, 0.35\ mm/rev$), and using one center point ($v = 60\ m/min\ and\ f = 0.25\ mm/rev$). Table (6.1) shows the design of the experiment .There are five runs which were done randomly. Each run was replicated at least 5 times. Each replication uses a new tool on which sequential inspections are performed in order to measure the wear. The wear is measured at discrete points of time through inspections using an Olympus SZ-X12 microscope. The procedure continues until the tool wear reached predefined threshold ($VB_{Bmax} = 0.2\ mm$). This procedure is repeated for 28 tools.

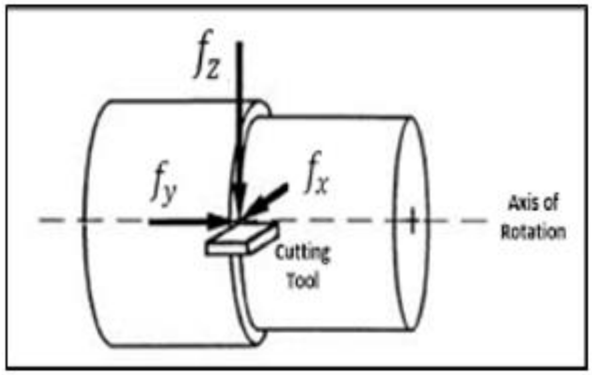


Figure 6-1: Forces directions during turning

Table 6.1: Design of experiment

Run	Cutting Speed (v) (m/min)	Feed Rate (f) (mm/ rev)	Depth of Cut (mm)	Number of replications
1	40	0.15	0.2	5
2	80	0.15	0.2	5
3	40	0.35	0.2	6
4	80	0.35	0.2	6
5	60	0.25	0.2	6

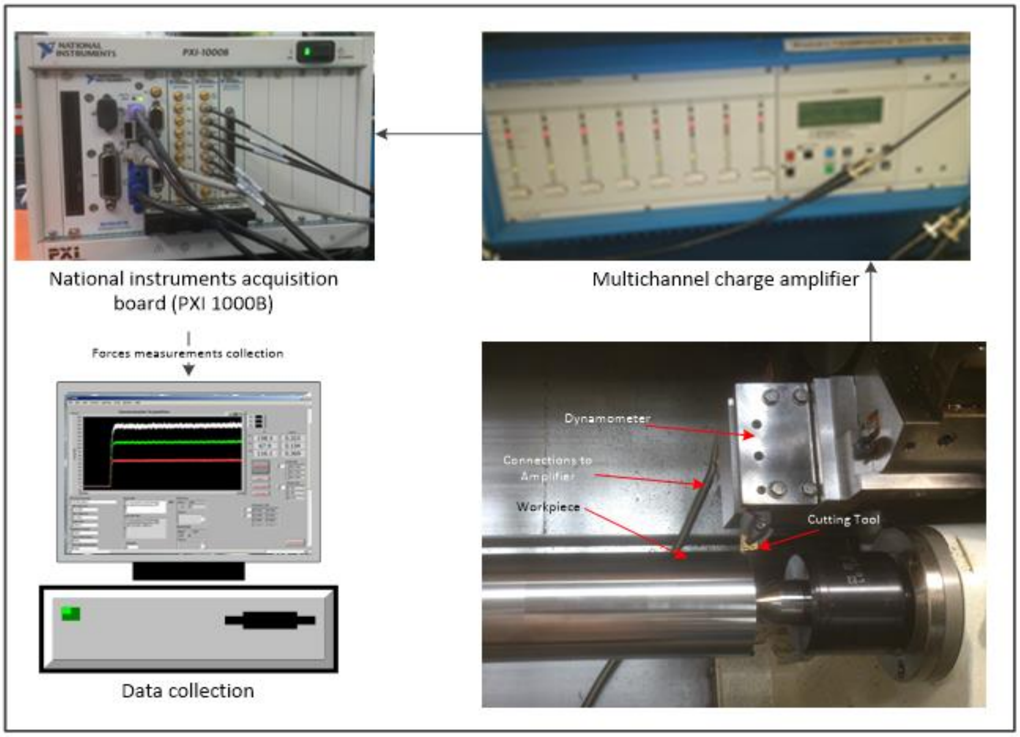


Figure 6-2: Schematic diagram of experimental setup

6.5 Knowledge extraction and learning

The cutting tool is failed when the tool is getting dull and no longer operates with acceptable quality (Gray et al. 1993). Predefining threshold on the maximum acceptable flank wear ($\mathrm{VB}_{\mathrm{Bmax}} = 0.2mm$) is a common way of quantifying the tool time to failure. The cutting tool fails after reaching the worn-out stage. The majority of publications classify tool wear only in two classes (Sick 2002). Two classes classification is all what we need if only fresh and significantly worn-out tools are our concern. Adjacent wear classes are defined by putting classification limit. In this work, we put the classification limit as ($\mathrm{VB}_{\mathrm{B}} = 0.15\,mm$). As such, 0.5 mm is kept as safety margin between fresh and worn-out stages. The same idea was taken by Rangwala (Rangwala 1988). He considered a classification limits to distinguish between fresh and worn-out tools. That classification limit is also chosen because, in some cases where the velocity is high, progressive wear is rapidly evolving and there is just one observation for wear value above 0.15 *mm* (i.e. before the tool fails).

Figure 6-3: Wear classification and failure threshold

Table 6.2: Experimental data of tool 1, replication 2

Inspection (observation) No $(\mathbf{t_i})$	Tool ID Run-replication	Time sec	Wear mm	Class {1,2}	Speed $(\boldsymbol{v})$ m/min	Feed $(\boldsymbol{f})$ mm/rev	Radial force $(\boldsymbol{f_x})$ N	Feed force $(\boldsymbol{f_y})$ N	Cutting force $(\boldsymbol{f_z})$ N
1	1-2	140	0.070	1	40	0.15	140.2	52.7	108.1
2	1-2	280	0.080	1	40	0.15	152.1	56.6	106.8
3	1-2	440	0.090	1	40	0.15	155.6	62.9	122
4	1-2	580	0.095	1	40	0.15	158.1	61.9	122.1
5	1-2	720	0.103	1	40	0.15	172.1	66.3	125
6	1-2	860	0.108	1	40	0.15	194	72.1	128.9
7	1-2	1000	0.113	1	40	0.15	215.1	75.9	139
8	1-2	1140	0.115	1	40	0.15	244.6	71.1	135.2

Table 6.3: Experimental data of tool 1, replication 2 (continued)

9	1-2	1280	0.120	1	40	0.15	268.5	77.7	143.4
10	1-2	1420	0.140	1	40	0.15	318.6	106.8	167.1
11	1-2	1492	0.145	1	40	0.15	319.9	84.1	160.7
12	1-2	1632	0.153	2	40	0.15	340.5	86.5	164.4
13	1-2	1772	0.160	2	40	0.15	354.2	88.7	167.5
14	1-2	1912	0.165	2	40	0.15	391	102.7	176.7
15	1-2	2112	0.21	2	40	0.15	467.10	103.40	169.80

Figure (6-3) shows the failure threshold and the classification limit. We have two classes: fresh tool (class 1), and worn-out tool (class 2). Data is collected for each replication (tool) and then classified into one of the two classes. For example, table (6.2) shows the classified experimental data of tool 1-2. This classification procedure is repeated for the 28 tools.

The software cbmLAD (c. Software 2012) is used, in order to extract the positive and negative patterns from the collected data, and then to train LAD to detect automatically the worn patterns, . The data from column 5 to column 10, in table (6.2), for 28 tools is used to find worn pattern using LAD technique. The variables are the external and internal covariates: cutting speed (v), feed rate (f), radial force (f_x), feed force (f_y), and cutting force (f_z), as shown in Table (6.3).

Table 6.4: Worn patterns

Worn patterns					
Pattern No	Speed (v) m/min	Feed (f) mm/rev	Radial force (f_x) N	Feed force (f_y) N	Cutting force (f_z) N
1	--	--	>416.65	>68.25	>164.30
2	--	--	>340.30, <428.55	>68.25	<184.85
3	<70	--	>364.65	>68.55	<221.60
4	<70	>0.2	>316.40	<97.10	<204.10
5	--	--	>293.35, <317.55	<82.95	<184.85

The data is also divided into two distinct spaces, one for fresh tool (class 1) when ($\mathrm{VB_B} < 0.15$ mm), and one for worn-out tool (class 2) when ($\mathrm{VB_B} \geq 0.15$ mm). Set O of the 273 observations is also divided into two sets of training, L, and testing, T. In this paper, tenfold cross validation procedure is conducted. As such, all the data (273 observations) is divided randomly into 10 sets of data, in which each class is represented in approximately the same proportion as in the full dataset. Each part is held out in turn, and the learning process is applied on the remaining nine-tenths of the data; then the quality of classification is calculated on the holdout (or testing) set. Thus, the learning procedure is repeated 10 times with different training sets. The results show that the quality of classification $\nu = 97.2$ %. Table (6.3) exhibits the worn patterns found by the software cbmLAD (c. Software 2012). The obtained five worn patterns are pure pattern. By pure, we mean that all five worn patterns don't cover any observation in fresh tool space. These patterns will lead us to build the on-line tool wear alarm system that is described in section 6.

6.6 The Statistical Proportional Hazards Model (PHM)

A Proportional Hazards Model (PHM) of the wear process is developed from the obtained experimental data. The alarm limit which is obtained from this model is compared to LAD on-line alarm system. Many researchers consider the PHM a good model for cutting tool life representation (Mazzuchi and Soyer 1989; V Makis 1995; Tail et al. 2010). In order develop the PHM, time to failure for each tool is required. We calculate the time to failure TTF by interpolating between two measurements around the failure threshold. TTF is calculated when tool wear threshold ($\mathrm{VB_{Bmax}} = 0.2\ mm$) is reached. For example, in Table (6.2), experimental results show wear evolution of tool 1-2 and by interpolating between the fourteenth inspection at ($\mathrm{t_{14}} = 1912\ sec, wear = 0.165\ mm$) and the fifteenth inspection at ($\mathrm{t_{15}} = 2112\mathrm{sec}, \mathrm{wear} = 0.21$ mm), the time to failure is found to be 2087 sec. This interpolation is repeated for the 28 tools. The results are given in Table (6.4).

Table 6.5: Times to failure for the 28 tools

Tool ID Run-replication	Time to failure sec	Tool ID Run-replication	Time to failure sec	Tool ID Run-replication	Time to failure sec	Tool ID Run-replication	Time to failure sec
1-1	1623.3	2-3	281.2	3-5	1230	4-6	102.5
1-2	2087	2-4	225.3	3-6	1006	5-1	233.5
1-3	1770	2-5	252.7	4-1	121.4	5-2	192
1-4	1524	3-1	1240	4-2	87.5	5-3	265
1-5	1560	3-2	1002	4-3	135	5-4	190
2-1	295	3-3	1320	4-4	135	5-5	160
2-2	267.5	3-4	1263.3	4-5	121.7	5-6	185

The concept of a PHM is that the failure rate of the cutting tool is not only dependent on the age of the tool, but is also affected by the internal and external covariates. As such, the failure rate consists of the product of a baseline failure rate $h_0(t)$, which is dependent only on the age of the tool, and a positive exponential function $\psi(Y,Z;\alpha,\gamma) = exp\{\sum_1^n \alpha_i Y_i + \sum_1^m \gamma_i Z_i\}$, where $n \, and \, m$ represent the number of external and internal covariates, respectively, $Y and \, Z$ represent the values of each external and internal covariate respectively, and $\alpha \, and \, \gamma$ represent the weight of each external and internal covariate. In general, the PHM is used to incorporate the internal and external covariates into the reliability modelling (H. Liu 1997; W. Wang 2004). The failure hazard rate at time (t) is expressed as in equation (6):

$$h(t,Y,Z;\alpha,\gamma) = h_0(t)\,\psi(Y,Z;\alpha,\gamma) \tag{6}$$

We consider the Weibull distribution as a baseline function. It is used extensively in modelling the tool failure (V Makis 1995; Tail et al. 2010; Mazzuchi and Soyer 1989). The failure hazard rate is written as:

$$h(t, Y, Z; \beta, \eta, \alpha, \gamma) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \ exp\{\sum_1^n \alpha_i Y_i + \sum_1^m \gamma_i Z_i\} \tag{7}$$

Where β is the shape parameter, η is scale parameter. The conditional survival function can thus be given as in equation (8).

$$R(t; Y, Z) = P(T > t|Y, Z) = \exp\left\{-\int_0^t h(t, Y, Z)dt\right\} \tag{8}$$

The conditional survival function $R(t; Y, Z)$, and its derivative $\dot{R}(t; Y, Z) = h(t; Y, Z)R(t; Y, Z)$ are used to estimate the parameters $(\beta, \eta, \alpha_1, \gamma_1, \gamma_2)$ by using the maximum likelihood function(Banjevic et al. 2001). EXAKT software estimates the PHM parameters as shown in figure (6-4).

Parameter	Estimate	Sign. (*)	Standard Error	Wald	DF	p - Value	Exp of Estimate	95 % CI	
								Lower	Upper
Scale	7.363e+004	-	5.932e+004	-	-	-	-	0	1.899e+005
Shape	2.727	N	0.9654	3.2	1	0.07363	-	0.8348	4.619
v	0.1415	Y	0.04841	8.545	1	0.003465	1.152	0.04662	0.2364
f	5.626	N	7.48	0.5657	1	0.452	277.5	-9.035	20.29
fx	0.02083	Y	0.003387	37.84	1	0	1.021	0.01419	0.02747
fy	0.004364	N	0.00935	0.2178	1	0.6407	1.004	-0.01396	0.02269
fz	-0.02678	Y	0.01334	4.034	1	0.0446	0.9736	-0.05292	-0.0006456

Figure 6-4: Summary of estimated parameters (based on ML method)

The column entitled "Sign." indicates whether the corresponding covariate was found to be significantly related to failure so the "Parameter" is significant (Y), or otherwise non-significant (NS). The cutting speed (v), the radial force (f_x), and the cutting force (f_z), are designated as significant (at this point in the analysis), while the feed (f), and the feed force (f_y) are not. Note that the feed force (f_y) has the lowest Wald Test result which represents the relative probability that feed force (f_y) has no significant impact on the risk of failure. The Wald Test is used to test if an independent variable has a statistically significant relationship with the risk of failure (dependant variable). The PHM model with all significant variables is found by eliminating the variables whose impact on the probability of failure is low. The final PHM model is shown in figure (6-5)

Parameter	Estimate	Sign. (*)	Standard Error	Wald	DF	p - Value	Exp of Estimate	95 % CI	
								Lower	Upper
Scale	1.294e+005	-	8.245e+004	-	-	-	-	0	2.91e+005
Shape	2.152	Y	0.5106	5.086	1	0.02412	-	1.151	3.152
v	0.1157	Y	0.02836	16.65	1	0	1.123	0.06012	0.1713
fx	0.02097	Y	0.003223	42.33	1	0	1.021	0.01465	0.02729
fz	-0.017	Y	0.006407	7.038	1	0.007981	0.9831	-0.02955	-0.004439

Figure 6-5: The model with only the significant variables (the best model to be used)

In the final PHM, the external covariate is the cutting speed (v). The internal covariates are the radial force (f_x), and the cutting force (f_z). The resulting hazard function is given in equation (9), where $n = 1, m = 2$.

$$h(t, Y, Z; \beta, \eta, \alpha_1, \gamma_1, \gamma_2) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\alpha_1 v+\gamma_1 f_x+\gamma_2 f_z} \tag{9}$$

$$h(t; Y, Z) = \frac{2.15}{129406}\left(\frac{t}{129406}\right)^{1.15} e^{0.1157v+0.02097f_x-0.017f_z}$$

The PHM concluded that the effects of the radial force and the cutting force are higher than the effect of the feed force on the progressive flank tool wear. This same conclusion was reached by the results of Huang and Liang in (Y. Huang and Liang 2005). EXAKT produces also the Kolmogorov-Smirnov test which evaluates the model fit. The summary of this goodness of fit test is shown in table (6.5).The test shows that the PHM offers a good modeling for the data.

Table 6.6: Summary of goodness of fit test results

Test	Observed value	P-value	PHM Fits Data
Kolmogorov-Smirnov	**0.200435**	**0.186612**	**Not rejected**

EXAKT gives a control-limit, $d > 0$, which is used in order to find the minimum expected machining cost per unit time (Aven and Bergman 1986; Viliam Makis and Jardine 1992). . The optimal stopping rule, in equation (10), is often used in condition based maintenance (CBM) as an alarm to when the tool should be replaced.

$$T_d = inf\{t \geq 0: Kh(t; Y, Z) \geq d\}, \tag{10}$$

where T_d is the preventive replacement time. The expected cost per unit time is expressed as:

$$\phi(T_d) = \frac{C\,P(T_d<T)+(C+K)\,P(T_d\geq T)}{W(d)} = \frac{C+K\,P(T_d\geq T)}{W(d)}, \tag{11}$$

where C is the preventive replacement cost, and $(C+K)$ is the failure replacement cost. $d^* = \phi(T_d^*)$ is the optimal cost at which the $\phi(T_d)$ is minimum, and T_d^* is the optimal time to replace the tool. $P(T_d \geq T)$ is the probability of a replacement due to failure ,and $P(T_d < T)$ is the probability of a preventive replacement. $W(d) = E(\min\{T_d, T\})$ is the expected replacement time. Optimal level d^* can be found by using the fixed-point iteration procedure (Banjevic et al. 2001; Viliam Makis and Jardine 1992), or by using Semi-Markovian Covariate Process (Bergman 1978). The warning function is derived when d^* is obtained. The warning function is derived from equation (10) as follow:

$$Kh(t, Y, Z) \geq d^* \tag{12}$$

$$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \mathrm{e}^{\alpha_1 v+\gamma_1 f_x+\gamma_2 f_z} \geq \frac{d^*}{K} \tag{13}$$

$$\alpha_1 v + \gamma_1 f_x + \gamma_2 f_z \geq \ln\left(\frac{d^*\,\eta^\beta}{K\beta}\right) - (\beta-1)\ln t \tag{14}$$

$$\alpha_1 v + \gamma_1 f_x + \gamma_2 f_z \geq \delta^* - (\beta-1)\ln t \tag{15}$$

$$Z^c \geq \mathrm{g(t)}, \tag{16}$$

where $\delta^* = \ln\left(\frac{d^*\,\eta^\beta}{K\beta}\right)$, $Z^c = \alpha_1 v + \gamma_1 f_x + \gamma_2 f_z$, and g(t)= $\delta^* - (\beta-1)\ln t$ is the warning function (Banjevic et al. 2001).

6.7 LAD on-line alarm system development and comparison with the PHM alarm function

The target is to build tool wear alarm system in order to detect the worn-out condition of the cutting tool, based on external and internal covariates' measurements. In figure (6-6), a schematic diagram shows how LAD on-line alarm system works. Off-line analysis is done in order to obtain the worn patterns as shown in section 4. In other words, the learning stage is done off-line. The worn patterns are then incorporated in LabVIEW software to build the alarm system. The operator defines the suitable cutting speed in the range of ($v = 40\ to\ 80\ m/min$) and the feed rate in the rangeof ($f = 0.15\ to\ 0.35\ mm/rev$). The cutting forces are measured during the turning process, and are transmitted to the alarm system. Upon the appearance of any worn pattern, the cutting tool should be replaced by the operator. Color-coded lamp instructs operator either to continue with the turning process, or to stop and change tool.

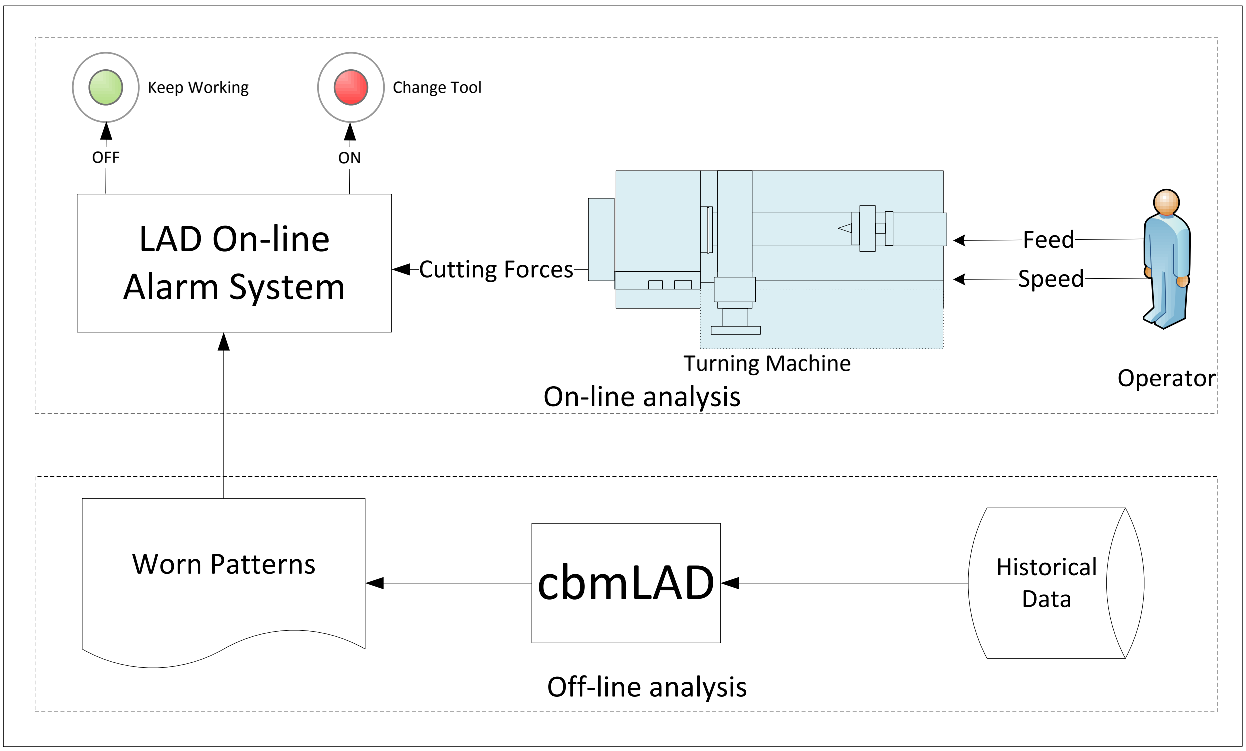


Figure 6-6: Schematic diagram for LAD on-line alarm system

The platform of PXI and LabVIEW were used to develop the on-line alarm system. The on-line alarm system detects the first appearance of a worn pattern, and gives alarm to the operator instructing him/her to stop the turning process and replace the cutting tool. The detected worn pattern number, and the tool working age are also indicated to the operator through labVIEW front panel. Color-coded lamp is incorporated to the front panel to alarm the operator when any worn pattern is detected. Additionally, the acquired forces data, the cutting tool working age, and the detected patterns are saved to the hard disk, with an automatically generated filename according to the time and date. Figure (6-7) shows a snapshot of the on-line alarm system front panel.

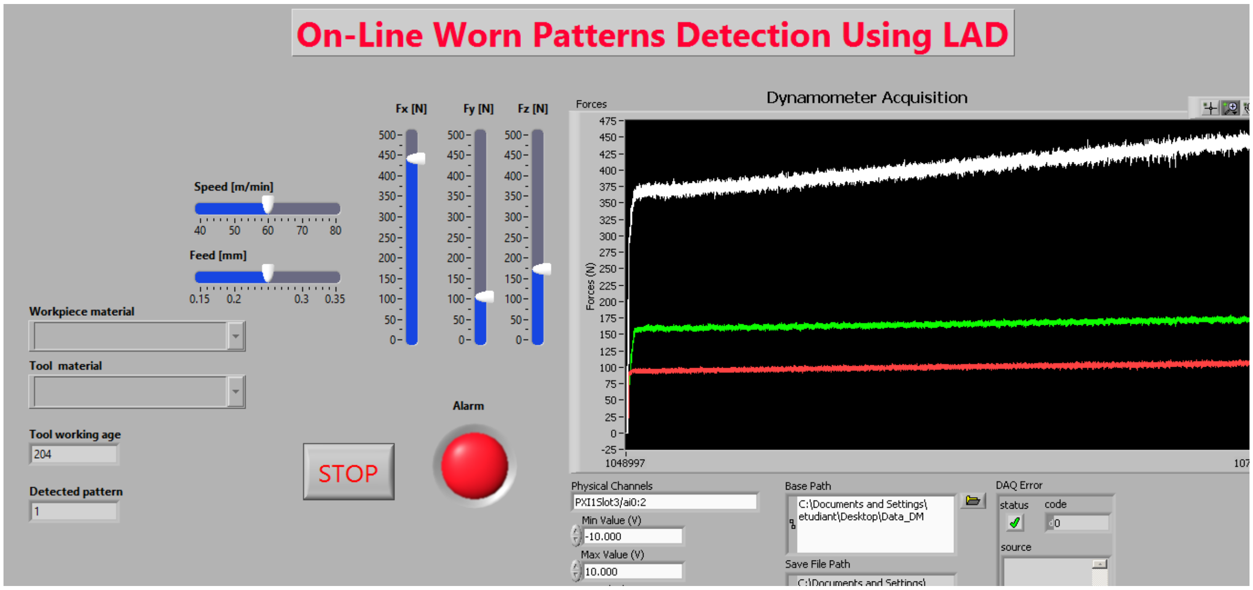


Figure 6-7: On-line alarm system front panel.

28 cutting tools (replications) were used to collect the data as discussed earlier and the data for each tool were stored. One replication from each of the 5 runs was chosen for testing the on-line alarm system. Replications (1-5), (2-5), (3-6), (4-6), and (5-6) were used for validation. For each replication, the covariates' measurements are transmitted to LAD' alarm system. The system has the worn patterns stored during the off-line learning stage. For each transmitted set of measurements the system search for worn patterns until color-coded lamp turns to red, when worn pattern is detected. Data report is automatically generated and saved. The report shows the tool working age, the detected worn pattern, and the corresponding covariates' values. At these values, the value of flank wear is listed in table (6.6). All data shows that the alarm system detects worn patterns before reaching the predefining threshold on the maximum flank wear ($\mathrm{VB_{Bmax}} = 0.2mm$), and around the classification limit of 0.15 mm.

Table 6.7: Replacement decision for PHM-model and LAD alarm

System's alarm readings						Validation			Comparison			
Tool ID Run-replication	Speed (v) m/min	Feed (f) mm/rev	Radial force (f_x) N	Feed force (f_y) N	Cutting force (f_z) N	LAD on-line alarm system			PHM- model Working age sec			TTF Experimental sec
						Working age sec	Wear mm	Worn Pattern NO	*r=2*	*r=1.5*	*r=1*	
1-5	40	0.15	401.2	68.3	145.2	1380	0.172	2	597	1020	937.7	1560
2-5	80	0.15	408.7	81.6	144.1	220	0.155	2	9	16	937.7	252.7
3-6	40	0.35	528.3	96.6	227.7	910	0.16	1	199	340	937.7	1006
4-6	80	0.35	522.4	139.6	243.2	90	0.15	1	5	9	937.7	102.5
5-6	60	0.25	368.1	91.3	97.5	135	0.153	2,3,4	72	123	937.7	185

Table (6.6) shows also the remplacement decision for the PHM based on its warning function. Each row has details of the detected worn pattern and corresponding values of machining conditions, cutting forces, cutting tool working, and flank wear value. In order to compare these results, the recommended optimal time to replacement is calculated by using the covariates' values in each row. The recommended optimal replacement time according to certain covariates' values using PHM (columns 10 and 11 in table (6.6)) are calculated using equation (17) which is derived from equation (16).

$$T_d^* = exp\left\{\frac{\delta^* - Z^c}{\beta - 1}\right\} \tag{17}$$

For example, in replication (1-5), if the cutting speed is ($v = 40$ m/min), the radial force is ($f_x = 401.2$ N), and the cutting force is ($f_z = 145.2$ N), Z^c is 10.57. According to equation (17), and considering a cost ratio r=2 (preventive replacement cost (C) is \$100, and the failure replacement cost ($C + K$) is \$200), T_d^* will be 597 sec., where $\delta^* = \ln(d^* \eta^\beta / K\beta) = 17.923$. Similarly, T_d^* =1020 sec when r=1.5 ($C = \$100,\ C + K = \$150$), and $\delta^* = 18.539$.

When r=1, the warning function g(t) is meaningless because $\delta^* = \infty$. As such, the expected replacement time $W(d)$ becomes the replacement time only at failure which is calculated from equation (11) as $W(d^*) = replacement\ cost/\phi(T_d^*) = replacement\ cost/d^*$ = 937.7 sec.

6.8 Discussion and Conclusion

In order to compare LAD's alarm system to the PHM statistical warning function, we compare between columns 7, 10, 11, 12, and 13. For example in replication (1-5), the tool's working age when the alarm was given is 1380 sec, the PHM warning time when r=1 is 937.7 sec, and the real experimental time to failure is 1560 sec. LAD used 88.46% of the tool life and the PHM used 38.26%when r=2, and used 65.38% when r=1.5. LAD is not affected by the value of cost ratio. In contrast, the statistical model will be more conservative when cost ratio is increasing because the statistical optimal time to replacement is decreasing. When the cost ratio is increasing, in order to minimize the cost per unit time, more frequent preventive replacement is recommended. When r=1, that is the cost effect is eliminated, the PHM recommended the run to failure, and in 3 cases out of 5 the tool failed before that time. This means that when we omit the effect of the cost ratio (putting r=1) and the warning function in the statistical PHM, LAD alarm system is still more accurate, in the sense that it is closer to reality.

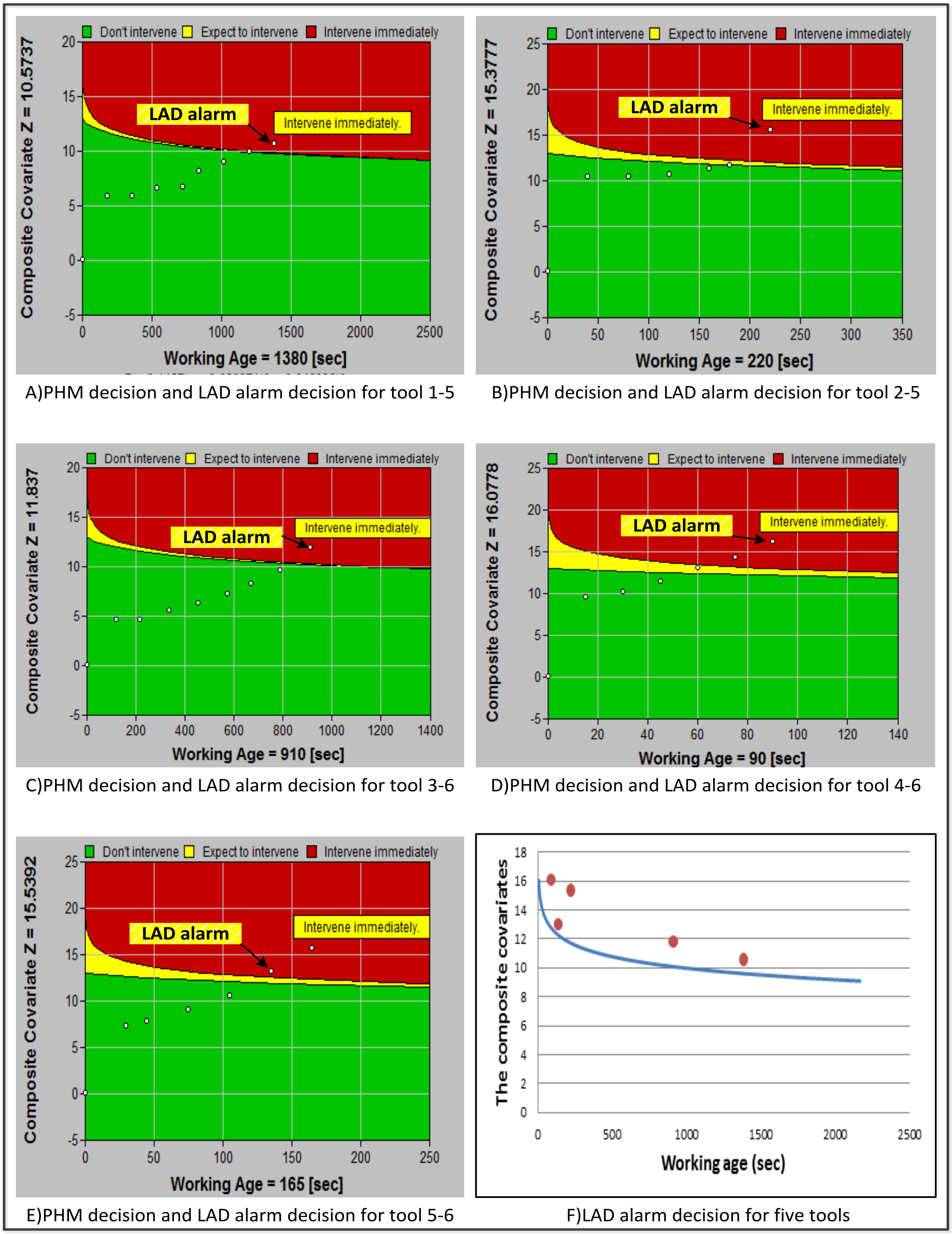
The tool working ages for the 5 replications are listed in table (6.6) for both the LAD alarm's system and the statistical PHM. In both LAD and PHM models, run (2-5) and run (4-6) have very low replacement time (columns, 7 10, and 11).This concludes that using carbides tool is not recommended when machining MMC at high speed, the same conclusion was also found before in (Kannan et al. 2006; Tomac et al. 1992). All five replications show that the statistical model is more conservative than LAD's alarm system based on pattern recognition. By conservative, we mean that PHM is replacing the cutting tool earlier than necessary, and consequently will cause loss of the valuable resource.

Moreover, we compare the five replications in EXAKT decision charts when r=2. Decision chart is built by plotting warning function. By defining the tool working age and the composite covariate $\boldsymbol{Z^c}$, the optimal decision is to determine whether the tool should be replaced immediately (the red area in figure (6-8)), or should be kept operating and be inspecting at the next inspection time (the green area), or should we keep operating but expect to replace before the next inspection time (the yellow area). Each replication was examined by using the previous experimental data in order to compare graphically between what the PHM decision model would have recommended for replacement tool and what LAD's alarm system would have recommended for replacement tool. The decisions are shown in figure (6-8). In figure (6-8-f), all LAD alarm decisions for the five tools are above the warning function $g(t)$ (the blue curve). This means that LAD gives an alarm to change the cutting tool at a higher working age than the statistical model. Obviously, in these cases, the LAD alarm system was capable of making the best action at the right moment. By best action we mean that replacement occurs before tool's failure and without losing much resource.

We recapitulate the discussion in the following: Statistical warning function $g(t)$ is plotted to differentiate between two decisions' areas the green area for 'keep working', and the red area for 'replace immediately'). Both decisions are based on the assumption of a statistical goodness of fit of a suitable hazard function, and the costs' ratio. LAD alarm points which are given in red in figure (6-8-f), are based on pattern recognition. As such, LAD replacement decision gave warning alarm before the tool wear reached the maximum flank wear $(VB_{Bmax} = 0.2mm)$ and without losing valuable resource due to early replacement. LAD can detect worn patterns on-line and in real time by monitoring covariates over time. In order to give accurate results, the only important requirement for using LAD is the availability of a database that represents accurately the phenomena under study. This is also a valid requirement for any statistical analysis and modelling.

In this paper, a new on-line tool wear alarm system based on LAD is developed. The alarm system is constructed based on data collected during turning titanium metal matrix composites (TiMMCs), under changeable machining conditions.The platform of PXI and LabVIEW were used to develop the alarm system. The LAD alarm system is validated by comparing it to the PHM warning function. The results show that the proposed alarm system detects the worn patterns and gives 'warning alarm' in order to replace the cutting tool at a working age that is relatively closer to the actual observed failure time.

In future work, the performance of the alarm system will be improved by including additional variables, such as vibration signal, acoustic emissions, and cutting temperatures. In order to distinguish between different tool wear phases, a multi-class LAD technique will be tested. The quality of the detected patterns will be improved, and non-pure patterns which can cover more than one class will be used, and give more details about the characteristics of LAD's patterns. Moreover, cbmLAD and our alarm system will be incorporated in a computer numerical control (CNC) machine; therefore, the learning stage can be done on-line thereby eliminating the need for off-line analysis.



A)PHM decision and LAD alarm decision for tool 1-5

B)PHM decision and LAD alarm decision for tool 2-5

C)PHM decision and LAD alarm decision for tool 3-6

D)PHM decision and LAD alarm decision for tool 4-6

E)PHM decision and LAD alarm decision for tool 5-6

F)LAD alarm decision for five tools

Figure 6-8: Replacement decision for PHM-model and LAD alarm (r=2)

CHAPTER 7 ARTICLE 5: TOOL REPLACEMENT BASED ON PATTERN RECOGNITION WITH LOGICAL ANALYSIS OF DATA

Yasser Shaban, Soumaya Yacout, Marek Balazinski

Accepted in:
The 61st Annual Reliability and Maintainability Symposium (RAMS2015)

7.1 Summary and conclusion

While traditional maintenance cost optimization is based on finding the reliability, and thus the probability of failure over time, in this paper, we show how to exploit condition monitoring data in machining operation in order to extract intelligent knowledge, and use this knowledge to determine the tool replacement time. This work is motivated by the increasing use of sensors in general, and specifically in condition monitoring. We show how the large volume of data that is now available in many industrial sites can give indications to the machining's operator in order to replace the tool. We use a methodology called Logical Analysis of Data (LAD). This methodology enables us to extract meaningful patterns that describe the state of the tool's wear, based on monitoring and measuring the cutting forces. Unlike the traditional experts' rule-based methods, the extracted patterns are not based on experts' opinion, but on information and hidden relations between the monitored forces.

We apply our methodology on data obtained from experiments that are conducted in the laboratory. The experimental data are collected during a turning process of titanium metal matrix composites (TiMMCs). These are new generation of materials which have proven to be viable in various industrial fields such as biomedical and aerospace, and they are very expensive.

In order to validate our methodology, we compare the results obtained when applying LAD to those obtained by using the well-known statistical Proportional Hazards Model (PHM). Findings and conclusion are given in the paper.

7.2 Introduction

TiMMCs have light weight and high strength which are suitable for aerospace industry to improve performance of aircraft. TiMMCs have high wear resistance, and they cause high wear rate on cutting tools. Progressive tool wear is the main cause of tool failure. Due to the fact that tool failure represents about 20% of machine down-time, and due to the high cost of machining, finding optimal tool replacement time is thus fundamental. The poor condition of the tool may leave negative effect on product quality in terms of dimensions, and surface finish. Therefore, product may be scrapped. When the material is very expensive such as Ti-MMCs, product scrapping causes an increase in the cost of machining. If the tool is replaced earlier than necessary, valuable resources will be lost (Tail et al. 2010).

Some researchers developed statistical methods based on pattern recognition to classify and monitor tool wear states. For example, (Damodarasamy and Raman 1993) developed a system for classifying tool wear states using pattern recognition. (Kang et al. 2007) developed a method of pattern recognition of tool wear based on discrete hidden Markov models. (S Purushothaman and Srinivasa 1994) developed a model for classifying a worn-out tool and a fresh tool using Artificial Neural Networks (ANN).

Using PHM in modeling cutting tool life is presented in (Shaban et al.). Many researchers consider internal covariates to model cutting tool life. Cutting forces, cutting temperatures, progressive wear, acoustic emissions and vibration signals are considered as internal covariates. For example, (H. Liu and Makis 1996) used PHM while considering the machining conditions as covariates, and they derived a formula to assess the tool reliability under variable cutting conditions. (P. H. Liu et al. 2001) used the PHM and stochastic dynamic programming for finding the optimal tool replacement times in a flexible manufacturing system. Tail et al (Tail et al. 2010) used a PHM to model the tool's reliability and hazard functions considering cutting speed as the model's covariate.

In this paper, the objective is to develop decision policy in cutting tool replacement using condition monitoring data in machining operation. We consider pattern-based technique using LAD, and we compare it to three statistical-based optimization models, namely cost optimization, availability optimization, and cost-availability optimization. In section 2, the experimental procedure is presented. In section 3, LAD is introduced. In section 4, the optimal replacement policies, and the decision rule are discussed.

7.3 Experiment description

The experiments were conducted in machining laboratory at École Polytechnique de Montréal. Experiment details are shown in table (7.1). Cutting forces signals are measured using 3-component dynamometer, which is connected to multichannel charge amplifier. The signals are then collected by national instruments acquisition board (PXI 1000B). In order to ensure adequate tool life, we limited the experiments to low cutting speed during turning metal matrix composites with carbides tools. The experiments were conducted, at low constant cutting speed $(\mathrm{v} = 40\mathrm{m/min})$, small constant depth of cut $(\mathrm{a_p} = 0.2\ \mathrm{mm})$, and two feed rate $(\mathrm{f} =$

$0.15\ \mathrm{mm/rev}$ and $f = 0.35\ \mathrm{mm/rev}$). At each feed rate, five experiments with five new tools were done. Figure (7-1) shows the experimental setup.

Table 7.1:Experiment details

Workpiece material	A cylindrical bar of Ti-6Al-4V alloy matrix reinforced with 10-12% volume fraction of TiC ceramic particles
Tool material	TiSiN-TiAlN nano-laminate PVD coated grades (Seco TH1000 coated carbide grades)
Equipment	A 6-axis Boehringer NG 200, CNC turning center

In figure 2, for each of the ten tools that are used, the wear and the forces are measured at sequential inspection points. At each inspection, the wear is measured by using an Olympus SZ-X12 microscope, and measured forces are recorded. For example, the experimental data of tool number 9, when $v = 40\mathrm{m/min}$ and $f = 0.35\ \mathrm{mm/rev}$ is shown in table (7.2), where 11 sequential inspection points are considered. This procedure continues until the tool wear reaches a predefined threshold of $VB_{\mathrm{Bmax}} = 0.2\ \mathrm{mm}$. This procedure is replicated for the ten tools. The collected data for the ten tools is shown in figure (7-2).

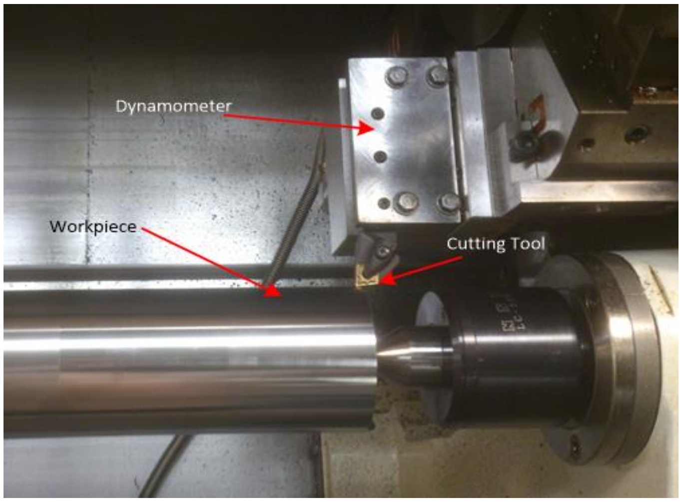


Figure 7-1: The experiment setup

In order to calculate the time to failure (TTF) when $\mathrm{VB_{Bmax}} = 0.2$ mm the wear evolution is interpolated linearly between the two values adjacent to $\mathrm{VB_{Bmax}} = 0.2$ mm. By interpolation, the TTF for each tool is calculated. For example, from Table (7.2), and by interpolating between the tenth and the eleventh inspections, TTF is found to be 1230 sec. This interpolation is repeated for the ten tools. The results for the ten tools are given in table (7.3).

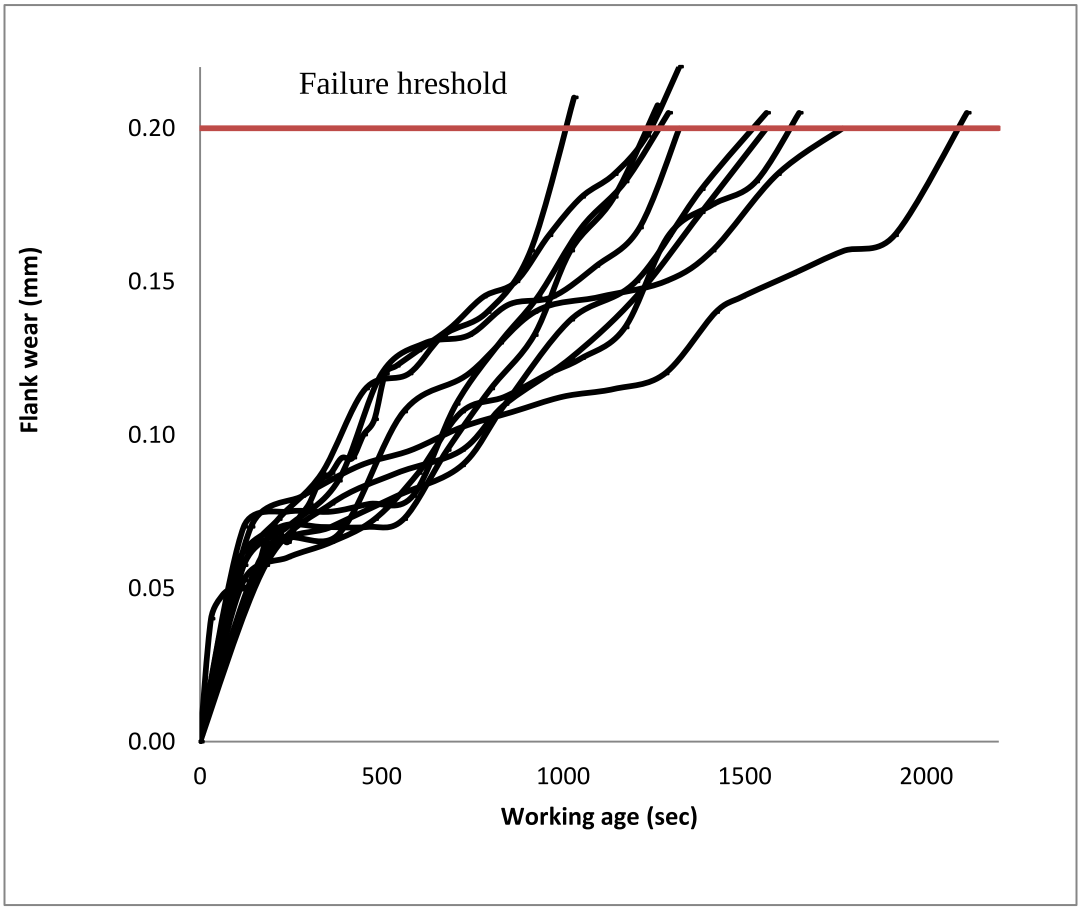


Figure 7-2: Tool wear measurements for 10 tools

Table 7.2: Experimental data of tool number 9

Inspection No	Time sec	Wear mm	Radial force (f_x) N	Feed force (f_y) N	Cutting force (f_z) N
1	100	0.0575	143.4	46.5	176.4
2	220	0.07	165	54.7	186
3	340	0.07	180	64	194
4	460	0.07	192.1	68	203
5	560	0.072	203.9	64.5	202.1
6	680	0.095	215.5	62.1	185.8
7	800	0.115	247.3	64.1	192.4
8	920	0.132	292.4	70.4	201.4
9	1020	0.16	342.5	68.7	203.8
10	1140	0.177	420.4	78.4	231.9
11	1260	0.207	485	94.2	247.7

Table 7.3: Times to failure for the 10 tools

$(v = 40m/min)$, $(f = 0.15\ mm/rev)$		$(v = 40m/min)$, $(f = 0.35\ mm/rev)$	
Tool ID	Time to failure sec	Tool ID	Time to failure sec
1	1623.3	6	1240
2	2087	7	1320
3	1770	8	1263.3
4	1524	9	1230
5	1560	10	1006

7.4 Logical Analysis of Data (LAD)

LAD is a combinatorial pattern recognition and classification technique. It is based on an artificial intelligence approach since it is applied in two consecutive stages, learning or training stage, and the testing stage. The main steps of the LAD are the binarization of data, the pattern generation, and the theory formation. Data binarization is the process of transformation of a data of any type into a Boolean database. The patterns generation procedure is the key building block in the LAD decision model. LAD is originally a two-class classification method, positive and negative class. A positive pattern is defined as a conjunction of some binary attributes, which is true for at least one positive observation and false for all negative observations in the training data set. A negative pattern is defined conversely. There are many techniques for pattern generation, for example enumeration, heuristics, and mixed linear programming (Ryoo and Jang 2009).Theory formation is the final step in the LAD decision model. A discriminant function is formulated in equation (1) in order to generate a score ranging between -1 and 1. When the output of a discriminant function is a positive value that means that the tested observation O belongs to the positive class, and negative otherwise. Zero value means no classification is possible (M.-A. Mortada et al. 2011).

$$\Delta(O) = \sum_{i=1}^{N^+} \alpha_i^+ Z_i^+(O) - \sum_{i=1}^{N^-} \alpha_i^- Z_i^-(O) \qquad (1)$$

where $N^+(N^-)$ is the number of positive (negative) patterns, $Z_i^+(O)\big(Z_i^-(O)\big)$ is equal to 1 if pattern (i) covers observation O, and is equal to zero otherwise, $\alpha_i^+(\alpha_i^-)$ is the weight of the positive (negative) pattern $p_i^+(p_i^-)$. For new observation O, the calculated value $\Delta(O)$ varies between +1 and -1, where +1(-1) is an indication of the domination of the positive (negative) patterns. ν, the quality of classification, is formulated in equation (2).

$$\nu = \frac{a+d}{2} + \frac{e+g}{4}, \qquad (2)$$

where a and d represent the proportion of observations, positive and negative, which are correctly classified respectively. e and g represent the proportion of observations, positive and negative, which are not classified respectively.

7.5 Optimal replacement & decision rule

In order to make an accurate comparison between LAD pattern recognition–based optimization, and the statistically-based optimization using a the PHM, we use the same wear states identification which is used by EXAKT software(Banjevic et al. 2001). This software develops a PHM. The tool wear is divided into five states which are used to construct the transition probability matrix. Each state represents particular stage in cutting tool wear progression. The wear states are considered as initial wear, slight wear which is regular state of wear, moderate wear which is micro breakage state of wear, severe wear which is fast wear state, and worn-out which is tool breakage. The flank wear bands are selected using the covariate distributions, and are not equally spaced. The last wear band, worn-out, does not contain many sample values (not more than 10%) which is considered reasonable as the highest state usually represents the most dangerous state. First band contains roughly 10% of the data values, second band 40%, third band 30%, and forth band 10%.

In pattern-based optimization, we reduce the problem into two-class which are normal- and critical class as shown in figure (7-3).

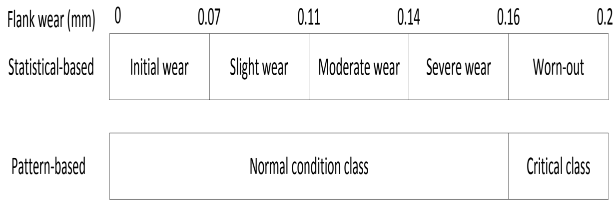


Figure 7-3: Tool wear identification on statistical-based and pattern-based

7.5.1 Pattern-based Optimization.

We used LAD to find the time to replacement T^*_{di} for each tool (i) when the wear reached worn-out state, that is the critical class. This class has special patterns which describe worn-out state. These patterns describe interactions between covariates which are feed rate (f), radial force (f_x), feed force (f_y), and cutting force (f_z) in critical class. The software cbmLAD is used to generate these patterns from the collected data. The data for the 10 tools contains 137 observations. This data is used to generate patterns that are representative of the two defined classes; the normal and the critical. The generated patterns for the critical class are Pattern 1, which is ($f_x > 341.6$, and $f_y > 68.3$, and $144.7 < f_z < 211$), Pattern 2 is ($f_z > 226.7$), Pattern 3 is ($364.6 < f_x < 414.9$). These patterns are detected only when the tool is worn-out, and they are never seen in the normal state. T^*_{di} is the time when any of these three patterns appears. It is recorded for each tool in table (7.4), with the corresponding values of covariates and the observed patterns' number. Obviously, the results show that T^*_{di} is different for each tool, since the wear propagation is a stochastic process. All the observations are divided into two sets of learning and testing. Then, tenfold cross validation procedure is conducted to validate and calculate the quality of classification. ν is found to be equal to 93.16 %. This means that in 93.16% of the observations, cbmLAD was able of defining correctly the state of the tool, either normal or worn-out, based on the detected patterns.

Table 7.4: replacement time for 10 tools based on LAD

Tool ID i	T^*_{di} *sec*	*Wear mm*	*Radial force* (f_x) *N*	*Feed force* (f_y) *N*	*Cutting force* (f_z) *N*	*critical Pattern No*
1	1290	0.1650	342.5	122.1	149.2	1
2	1772	0.1600	354.2	88.7	167.5	2
3	1410	0.1600	387.1	108.3	190.2	1,3
4	1380	0.1800	423.8	77.3	150.5	1
5	1380	0.1725	401.2	68.2	145.2	3
6	1050	0.1775	381.3	109.7	209.1	1,3
7	1210	0.1675	540.5	170.9	249.1	2
8	1050	0.1675	367.3	81.6	216.7	3
9	1020	0.1600	342.5	68.7	203.8	1
10	910	0.1600	528.3	96.6	227.7	2

7.5.2 Statistical-based Optimization.

PHM is used in order to model the experimental data. The concept of PHM is that the failure rate of cutting tool depends on the age of the tool and covariates. The failure rate is represented as the product of a baseline failure, which depends on the age of the tool, and a positive function, $\exp\{\sum_1^m \gamma_i Z_i\}$, that represents the effect of covariates on failure rate. Where m represents the number of covariates, Z represents the value of each covariate, and γ represents the weight of each covariate. In this work, the baseline is considered Weibull hazard function. The failure hazard rate at time (t) is given as follows:

$$h(t, Z; \beta, \eta, \gamma) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\{\textstyle\sum_1^m \gamma_i Z_i\} \qquad (3)$$

Where β is the shape parameter, η is scale parameter, the conditional survival function is given as in equation (4),

$$R(t;Z) = \exp\{-\left(\frac{t}{\eta}\right)^{\beta} \exp\{\sum_{1}^{m} \gamma_i Z_i\} \quad (4)$$

The conditional survival function $R(t;Z)$ and its derivative $\dot{R}(t;Z) = h(t,Z(t))R(t;Z)$ are used to estimate the parameters (β, η, γ) by using maximum likelihood function (Banjevic et al. 2001). EXAKT software is used to estimate the PHM model and its parameters. The covariates are feed rate (f), radial force (f_x), feed force (f_y) and cutting force (f_z). All combinations are examined and tested. The best model with significant variables is found by eliminating the variables whose impact on the probability of failure is low and making comparison between obtained models. The best model is showed as in equation (5) when radial force is taken as the model's covariate.

$$h(t,Z) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma f_x} = \frac{5.99}{7406}\left(\frac{t}{7406}\right)^{4.99} e^{0.022\, f_x} \quad (5)$$

Kolmogorov-Smirnov test (K-S test) evaluates the model fit. The test shows that the PHM offers a good modeling for the data with P-value equal to 0.889842.

7.5.2.1 Cost Optimization

The cost function is given as follows:

$$\phi(T_d) = \frac{C_p\, P(T_d < T) + C_f\, P(T_d \geq T)}{W(d)} \quad (6)$$

where T is the failure time, T_d is the preventive replacement time, C_f is the failure replacement cost, and C_p is the preventive replacement cost. The optimal cost is achieved when (T_d) is minimum , and where T_d^* is the optimal time to replace. $P(T_d < T)$ is the probability of preventive replacement, $P(T_d \geq T)$ is the probability of failure replacement, and $W(d) = E(\min\{T_d, T\})$ is the expected replacement time. $C_p = \$100$, and $C_f = \$200$, and (K) is the difference between replacement costs includes the extra costs due to the consequences of failure replacement. The cost analysis gives the minimum cost ($d^* = \$0.0816$ /sec) associated with the optimal risk level to intervene ($Kh = \$0.08163$/sec). In order to minimize cost, cutting tool should be replaced at $T_d^* = 1348.33$ sec when the optimal hazard rate is $h = 8.16 \times 10^{-4}$ /sec.

7.5.2.2 Availability Optimization

The availability function is as follows:

$$A(T_d) = \frac{uptime}{uptime + downtime}$$

$$= \frac{W(d)}{W(d)+T_p\, P(T_d<T)+T_f\, P(T_d \geq T)} \qquad (7)$$

Availability (A) is the percentage of time that cutting tool is available for machining. The optimal availability is achieved when $A(T_d)$ is maximum, where T_d^* is the optimal time to replacement, $T_p = 160$ sec is the preventive replacement time, and $T_f = 540$ sec is the failure replacement time. K is the difference between replacement times that includes the extra time due to the consequences of failure replacement. The availability analysis shows the optimal risk level to intervene at ($Kh = 0.14169$). In order to maximize availability, cutting tool should be replaced at $T_d^* = 1249.4$ sec when the optimal hazard rate is $h = 3.72 \times 10^{-4}$ /sec.

7.5.2.3 Cost-availability Optimization

cost-availability optimization is the combination of cost and availability optimization. We minimize the cost per unit time while taking into our consideration replacement costs, replacement times and costs of downtimes. For example, $\text{total preventive replacement cost} = C_p + a \times T_p$, where C_p is the fixed cost, a is the cost per unit time, and T_p is the down time.

$$\psi(T_d) = \frac{(C_p+a\,T_p)\,P(T_d<T)+(C_f+a\,T_f)\,P(T_d \geq T)}{W(d)+T_p\,P(T_d<T)+T_f P(T_d \geq T)} \qquad (8)$$

The values of replacement costs in cost analysis ($C_p = \$100, C_f = \$200$), the values of replacement times in availability analysis ($T_p = 160$ sec, $T_f = 540$ sec), and the cost per unit time during downtimes ($a = \$0.15$ /sec) are used to find the optimal replacement time. The cost and availability analysis shows the optimal risk level to intervene ($Kh = \$0.1087$ /sec). In order to maximize availability and minimize cost simultaneously, cutting tool should be replaced at $T_d^* = 1339.52$ sec when the optimal hazard rate is $h = 7.6 \times 10^{-4}$ /sec.

7.6 Decision Rule

The decision rule gives us the optimal time to replacement when considering cost and/or availability analysis, and the decision of tool replacement or to keep working and monitoring the covariate (radial force) at discrete time intervals (Banjevic et al. 2001). The optimal replacement decision is shown as in equation (9),

$$T_d^* = inf\ \{t \geq 0: Kh(t, Z(t)) \geq d^*\} \qquad (9)$$

Where ($d > 0$) is a control-limit value and d^* is the optimal level which was found in section 4.2. K is the difference between replacement costs C_f and C_p in cost analysis, the difference between replacement times T_f and T_p in availability analysis, and the difference between total replacement costs including downtime costs in the combination of cost and availability analysis. The warning level function which helps to make replacement decision is derived.

$$\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{\gamma Z} \geq \frac{d^*}{K} \qquad (10)$$

$$\gamma Z \geq ln\left(\frac{d^*\,\eta^{\beta}}{K\beta}\right) - (\beta - 1)\ ln\, t \qquad (11)$$

The function $g(t) = \ln\left(d^*\,\eta^{\beta}/K\beta\right) - (\beta - 1)\ln t$ is consider as warning level function, applied to an overall covariate value $Z^c(t) = \gamma Z$. The warning level function $g(t)$ is built for each policy. By monitoring the radial force at discrete time intervals, the composite covariate $Z^c(t) = \gamma Z = 0.022\ f_x$ is calculated, and by defining the tool working age, $g(t)$ is calculated. As such, we can make a decision that will optimize the long-run maintenance cost and/or availability. In figure (7-4), the optimal decision is to determine whether the tool should be replaced immediately, if it information point is above the curve, or should we keep operating and inspect at the next inspection time. This case is represented by all the points that are under the curve. The three curves in figure (7-4) represent the three optimization functions of cost, availability, and cost-availability.

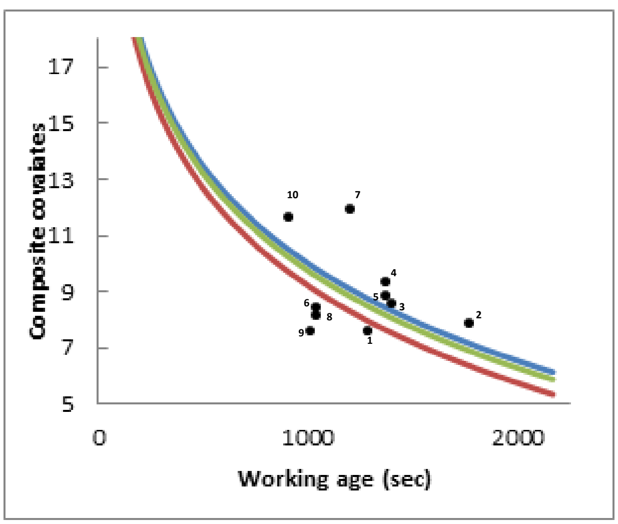


Figure 7-4: Replacement decisions

To summarize, the optimal policy suggests replacement at time (t) for which $0.022\ f_x \geq 44.47 - 4.99 \ln t$ in the cost analysis (blue curve), and replacement at t for which $0.022\ f_x \geq 43.68 - 4.99 \ln t$ in the availability analysis (red curve), and suggests replacement at t for which $0.022\ f_x \geq 44.21 - 4.99 \ln t$ in cost-availability analysis (green curve). In contrast, in pattern-based technique, optimal decision is made when the tool wear start the worn-out state, the critical state, and a worn-out pattern is observed in black points. At this moment, the tool should be replaced immediately. For example, replacement decisions for the ten tools are shown in table (7.5) and figure (7-4). In most cases, 6 out of 10, pattern-based technique recommends replacement later than statistical techniques, that is above the curves, and closer to the actual failure. This means that pattern-based technique saves tool's valuable resource. For example, tool number 7, the statistical -based suggests replacement at 684.6 sec, 584.3 sec, and 649.8 sec in cost analysis, availability analysis, and in cost-availability analysis respectively, while pattern–based analysis suggests replacement at 1210 sec and the tool fails at 1320 sec. For the remaining 4 tools under the curves, statistical techniques recommends "keep working" while the tool wear reaches the critical state, which means that the pattern-based technique is still accurate, in the sense that it is close to reality. For example, tool number 9, the statistical optimal policy suggests replacement after failure while pattern–based analysis suggests replacement before failure. This paper presents a pattern-based technique using LAD in order to blaze a new trail in optimal replacement techniques.

Table 7.5: replacement times comparison for ten tools

Tool ID i	Replacement Times-based (sec)				TTF sec
	Cost	Availability	Cost - availability	Pattern-based	
1	1638.9	1398.9	1555.7	1290	1623.3
2	1556.5	1328.6	1477.5	1772	2087
3	1346.3	1149.2	1278.0	1410	1770
4	1145.2	977.5	1087.0	1380	1524
5	1265.2	1079.9	1200.9	1380	1560
6	1381.2	1179.0	1311.1	1050	1240
7	684.6	584.3	649.8	1210	1320
8	1469.1	1254.0	1394.5	1050	1263.3
9	1638.9	1398.9	1555.7	1020	1230
10	722.4	616.6	685.7	910	1006

CHAPTER 8 ARTICLE 6: CUTTING TOOL WEAR DETECTION USING MULTI-CLASS LOGICAL ANALYSIS OF DATA

Yasser Shaban, Soumaya Yacout, Marek Balazinski, Krzysztof Jemielniak

Submitted to:

Journal of intelligent manufacturing

8.1 Abstract

This paper presents a new tool wear multi-class detection method. Based on experimental data, tool wear classes are defined using Douglas-Peucker algorithm. Logical Analysis of Data (LAD) is then used as a machine learning, pattern recognition technique for double objective of detecting the present tool wear class based on the recent sensors' readings of the time dependant machining variables, and deriving novel information about the role of machining variables by doing patterns analysis. LAD is data driven technique which relies on combinatorial optimization and pattern recognition. The accuracy of LAD is compared to Artificial Neural Network (ANN), since ANN is the most familiar machine learning technique. The proposed method is applied on experimental data which are gathered under various machining conditions. The results show that the proposed method detects the tool wear class correctly and with high accuracy.

Keywords

Tools wear detection, logical analysis of data, pattern recognition.

8.2 Introduction

In the scientific literature, the tool wear detection in machining processes is estimated by two approaches: Continuous tool wear estimation and tool wear classification (Sick 2002). Researchers estimate tool wear continuously using data driven techniques e.g. (Marek Balazinski et al. 2002; Achiche et al. 2002; Ren et al. 2008). (Sick 2002) presented an exclusive review of online tool wear detection in turning and found that the majority of researches considered tool wear detection based on tool wear classification. For example, (Damodarasamy and Raman 1993) used three adjacent classes of tool wear in order to develop a detection tool wear model using pattern recognition. They concluded that pattern recognition can be used to detect the classes of cutting tool wear. (S Purushothaman and Srinivasa 1994) developed tool wear monitoring model by using two classes, worn-out tool and a fresh tool. They used Artificial Neural Networks (ANN) for building the model. (Ertunc and Oysu 2004) used five classes to develop tool wear monitoring system using dynamic hidden Markov models. (Kang et al. 2007) developed tool wear monitoring model using pattern recognition based on discrete hidden Markov models. A three classes cutting tool wear model is used. (Tobon-Mejia et al. 2012) used five classes in order to diagnose the wear's progression. They used dynamic Bayesian networks' technique.

Although the continuous tool wear estimation gives a better picture of the tool wear progression, many researchers consider that in practical situations the tool wear classification is quite sufficient for allowing the operator to make an informed decision (Sick 2002). In this paper we present a novel tool wear classification approach. Since all the previously mentioned methods for tool wear detection and classification share one common disadvantage, which is the use of statistical techniques that impose some statistical assumptions, such as the prior distribution, or the independence of machining variables. In this paper, we introduce a new technique that doesn't rely on any statistical assumptions. The technique is called Logical Analysis of Data (LAD), which is a machine learning pattern recognition data-driven approach. As many machine learning approach, LAD is applied in two steps; first the training step at which LAD learns about the wear process from the experimental results, then the application step in which LAD detect the wear class based on the learning acquired in the previous step. Since LAD is a supervised learning technique, the experimental results are characterized by their corresponding classes. In this paper, this is accomplished by using the Douglas-Peucker (DP) algorithm(Douglas and Peucker 1973).

The progressive wear is monotonically increasing; therefore, it's very important to define adjacent wear classes in order to see how wear classes are actually considered when machining process are carried out (Sick 2002). Another importance of wear classification that flank wear limit which is always used as indication for cutting tool life is not the same when cutting tool is used in rough and finish cut (Damodarasamy and Raman 1993). For example, if we have 5 classes of wear, the first and second classes can represent the wear limit for finish cuts while third and fourth classes can represent the wear limit for rough cut, and finally the fifth class can represent the tool failure.

This paper is organized as follows: In section 8.3, the methodology of LAD is introduced. The experimental procedure and wear classification method are presented in section 8.4. In section 8.5, results and discussion are introduced. Concluding remarks are given in section 8.6.

8.3 Logical analysis of data (LAD)

LAD is knowledge discovery technique first developed by Peter Hammer in 1986 (Peter L Hammer 1986; Crama et al. 1988). LAD allows the classification of phenomena based on pattern recognition. To generate patterns, two consecutive phases are applied, training or learning phase, and the testing or the theory formation phase. In the learning phase, part of the dataset is used to extract hidden patterns. In testing phase, the remainder of the dataset is used to test the accuracy of the previously learned knowledge from the extracted patterns. The methodology of LAD is composed of three steps: Data binarization, pattern generation, and theory formation.

Data binarization

In data binarization, the dataset is transformed into a Boolean dataset. The binarization step involves the transformation of variables' values to binary attributes using a binarization technique which is discussed in (Bores et al. 2000). Here, the variables are the machining conditions. For example, we consider that A is a continuous numerical variable. To transform A into binary attributes, we start by ranking the numerical values, u, of A in ascending order, as follows:

$$u_A^{(1)} < u_A^{(2)} < \cdots < u_A^{(Q)} \qquad (Q \leq N) \qquad (1)$$

Where Q the total number of is distinct values of the variable A and N is the total number of observations in the training set. The cut-point $\varepsilon_{A,j}$ is found between each pair of values that belong to different classes. The cut-point is calculated by averaging these two values as shown in equation (2):

$$\varepsilon_{A,j} = (u_A^{(k)} + u_A^{(k+1)})/2 \qquad (2)$$

Where $u_A^{(k)}$ and $u_A^{(k+1)}$ belong to different classes. A binary attribute b is then formed from each cut-point. Each cut-point $\varepsilon_{A,j}$ has a corresponding binary attribute $b_{\varepsilon_{A,j}}$ with defined values as in equation (3). The number of binary attributes that represent the continuous variable A is thus equal to the number of cut-points.

$$b_{\varepsilon_{A,j}} = \begin{cases} 1 & \text{if} \quad u_A \geq \varepsilon_{A,j} \\ 0 & \text{if} \quad u_A < \varepsilon_{A,j} \end{cases} \tag{3}$$

Pattern generation

Pattern is hidden rules that describe one class of wear and not the others. Here, the class refer to a specific cutting tool wear stage. A pattern is defined as a conjunction of literals which is true for at least one observation in a specific class and false for all observations in other classes in the training dataset. A literal is a Boolean variable x or its negation $\bar{x}$ (Bores et al. 2000). b_i is a binary attribute in the training set and can be represented by a literal x_i or its negation $\bar{x}_i$. If $b_i = 1$ then x_i is true and if $b_i = 0$ then x_i is false. Similarly, $\bar{x}_i$ is true when $b_i = 0$ and false when $b_i = 1$. The number of literals used to define the pattern is called the degree of a pattern d. Pattern p of degree d is a conjunction of d literals. For example, consider a binarized data set consisting of five binary attributes $(b_1, b_2, b_3, b_4, b_5)$. A conjunction of literals $\bar{x}_1, x_2, x_4, \bar{x}_5$ is said to be a pattern of degree 4 in a specific class if at least one observation in that class has the respective values (0,1,1,0) for attributes (b_1, b_2, b_4, b_5) and no observation in all the other classes has these values. A pattern covers a certain observation in the training set if and only if it is true for that particular observation (Bores et al. 2000).

There are many techniques for pattern generation such as enumeration (Bores et al. 2000), heuristics (Peter L Hammer 1986; P.L. Hammer and Bonates 2006), and linear programming (Ryoo and Jang 2009). In this paper we follow the pattern generation technique which has been proposed in (Ryoo and Jang 2009). In that paper the pattern generation problem is converted to a set covering problem, and then solved by a mixed integer linear programming (MILP) technique. It should be noted that using that technique to generate patterns does not mean that LAD based decision model is linear. In the MILP problem, it is assumed that each generated pattern p is associated with a Boolean pattern vector $W = (w_1, w_2, \dots w_q, w_{q+1}, w_{q+2}, \dots w_{2q})$ with size n, where $n = 2q$, and q is the size of the binary observation vector. The elements of the Boolean pattern vector W is restricted to the following conditions:

$$w_j + w_{j+q} \leq 1 \qquad \forall\, j = 1,2, \ldots q \tag{4}$$

In other words, if $w_j = 1$ then the literal x_j is included in pattern p. Similarly, if $w_{j+q} = 1$ then literal $\bar{x}_j$ is included in pattern p. A pattern p cannot include both the literal x_j and its negation $\bar{x}_j$ at the same time. The pattern p can be deduced after getting the Boolean pattern vector W as a solution of the set-covering problem. For example, if the Boolean pattern vector $W = (0,1,0,0,0,0,0,0,0,1)$, it means that $p = x_2\bar{x}_5$, $q = 5$, and The pattern degree $d = 2$.

For the generation of a pattern p^m which belongs to class c_m, where m is the class number, $Y = (y_1, y_2, \ldots y_{D^m})$) is the Boolean coverage vector whose number of elements equal to the number of observation D^m in class c_m, and where y_i is equal to 0 if a pattern p^m covers the observation i and 1 otherwise in c_m. Minimizing Y means finding pattern which belongs to c_m that covers the maximum number of observations of this class. Each observation $i \in \pi^m$ is represented as a Boolean observation vector $a_i = (a_{i,1}, a_{i,2}, \ldots a_{i,q}, a_{i,q+1}, a_{i,q+2}, \ldots a_{i,2q})$ such that $a_{i,j} = 1$ if the binary attribute $b_j = 1$, and $a_{i,q+j} = 1$ if $b_j = 0$, and π^m is the set of observations in class *m*.

The MILP objective is to minimize Y. This means to minimize the number of observations in c_m , which are not covered, since each observation in the training dataset must be covered by at least one pattern in order to differentiate the observations of each class. In this optimization problem, the decision variables are the pattern vector W, the degree d, and the coverage vector Y. The constraints for that optimization problem are as follows:

By definition, a pattern p^m must not cover any observations in any class other than c_m . For that reason, the dot product of the pattern vector W and the observation $i \in \Pi^-$ must be less than the degree d of the pattern p, where Π^- is the set of all the observations in the training dataset that are not in c_m :

$$\sum_{j=1}^{2q} a_{i,j}\, w_j \le d - 1 \qquad \forall\, i \in \Pi^- \qquad (5)$$

The generated pattern should cover at least one observation $i \in \pi^m$, but not necessarily all the observations in π^m. This condition can be formulated for each observation as:

$$\sum_{j=1}^{2q} a_{i,j}\, w_j + q y_i \ge d \qquad \forall\, i \in \pi^m \qquad (6)$$

The newly-generated pattern must not be a subset of any of the patterns that have already been generated. Every generated pattern vector W is stored as vector v in the set V_m containing all pattern vectors of the patterns generated previously for the set c_m:

$$\sum_{j=1}^{2q} v_{k,j}\, w_j \le d_k - 1 \qquad \forall\, v_k \in V_m \qquad (7)$$

All previous steps can be summarized as follows:

The objective function

$$\min_{W,Y,d} \sum_{i \in \pi^m} y_i$$

$$\text{s.t} \begin{cases} (4), (5), (6), (7) & \\ \sum_{j=1}^{2q} a_{i,j}\, w_j = d \quad \forall\, i \in \pi^m & (8) \\ 1 \le d \le q & (9) \\ W \in \{0,1\}^{2q} & (10) \\ Y \in \{0,1\}^{D^m} & (11) \end{cases}$$

As a result, the linear set-covering problem mentioned previously generates the strongest pattern. A strongest pattern covers a maximum number of observation which others pattern can’t do. An iterative mechanism is needed to generate an entire set of patterns belongs to c_m. The same procedure is repeated in order to generate the patterns which belong to other classes.

Theory formation

The discriminant function used in the multi-class LAD approach different from the classical one which is used in two-class LAD decision model. In the two-class LAD approach, the output of discriminant function is between -1 and 1. Positive output means the tested observation belongs to the positive class, and negative otherwise. Zero value means no classification is possible (M.-A. Mortada et al. 2011). Here, in multi-class, output sign of a discriminant function is no longer sufficient to classify a new unknown observation. A single separator between class c_m and all other classes is built, As such, Z different two-class classifiers are built, where Z is the total number of classes. Let S_m be the mth classifier separating observations in class c_m and observations in Π^-, not in c_m. For each new observation O the calculated value $\Delta(O)$ is given a score for each class. Therefore, a new observation belongs to the class with the highest score. If $\Delta(O)$ has the same value for two or more different classes, then the observation O is unclassified.

Discriminant function $\Delta(O)$ is formulated as following (Herrera and Subasi) :

$$\Delta(O) = \underset{m=1,\ldots,Z}{arg\ max}\, S_m(O) \qquad (12)$$

Where $S_m(O) = \sum_{P_n \in \mu_m} \alpha_n P_n(O)$, P_n is corresponding pattern in set μ_m covers observation O, and $\alpha_n > 0$ is the coverage rate of that pattern $P_n \in \mu_m$ $(m = 1, \ldots, Z)$ with respect to the observations of class. Pattern $P_n(O) = 1$ if it covers observation O and zero otherwise.

The accuracy (AC) of classification is

$$AC = \frac{\sum_m^Z G_m}{\mathrm{Nt}},$$

where G_m is the total number of correctly classified observations in class m, and Nt is the total number of observations in the testing set.

8.4 Experiment description and wear classification

The experiments were conducted on a conventional lathe TUD-50. A CSRPR 2525 tool holder equipped with a TiN–AL2O3–TiCN coated sintered carbide insert SNUN 120408 was used in the tests. AISI 1045 steel was used as workpiece material. A six cuts with different cutting variables were conducted sequentially in order to remove six parts as shown in figure (8-1). The values of the machining controllable variables are shown in figure (8-1), where f is the feed rate in $\mathrm{mm/rev}$, and t is the machining time. The feed force (F_f), and cutting force (F_C) are the uncontrollable machining variables which have values that depend on the wear. They were measured by using a Kistler 9263 dynamometer. The cutting speed was selected to ensure approximately the same share in tool wear. The two tools that were used had a soft, cobalt-enriched layer of substrate under the coating. Their tool life ends after this coating wore through. Two identical tools were used. With tool 1, 10 cycles, of six cuts each, were performed until the flank wear $\mathrm{VB_B}$ reached approximately 0.5 mm. After each cycle, the tool wear was measured and its value corresponding to single cut was linearly interpolated. The collected data contains 71 observations for tool 1. For example, table (8.1) shows the collected data of cut 5 in tool 1 experiment. In the tool 2, failure of the coating resulted in chipping of the cutting edge at the end of 9th cycle. At this moment, flank wear was about 0.35 mm. The collected data for tool 2 contains 66 observations. Figure (8-3) shows the feed and the cutting forces, vs. the tool wear in tool 1 and 2. The legend in each graph indicates the feed (f) in mm/rev and the depth of cut ($\mathrm{a_p}$) in mm for each cut. The feed force (F_f) is independent of feed rate (f), but being affected by the depth of cut ($\mathrm{a_p}$) and the tool wear. The cutting force (F_c) depends on depth of cut ($\mathrm{a_p}$) and feed rate (f), while being weakly affected by tool wear. As such, we don't need information about the depth of cut ($\mathrm{a_p}$) to define the tool wear class. Here, tool wear class is defined by using three machining variables: feed rate (f), cutting force (F_c), and feed force (F_f).

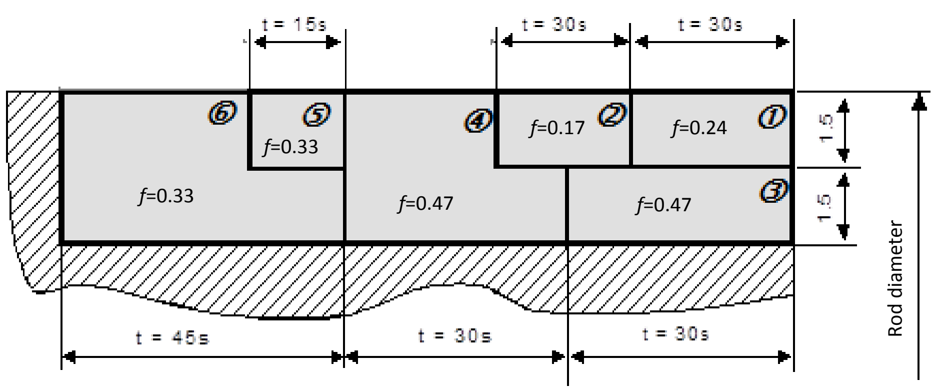


Figure 8-1: Six cuts with different cutting variables

As discussed in (Ertunc and Oysu 2004; Tobon-Mejia et al. 2012), the tool wear is classified into five classes. Each class represents particular state in cutting tool wear progression. The five wear classes are: (1) the initial wear, (2) the slight wear which is the regular state of wear, (3) the moderate wear which is the micro breakage state of wear, (4) the severe wear which is the fast wear state, and (5) the worn-out which is tool breakage state. LAD is based on supervised learning; this means that all the observations should be labeled by a class number before analysis. Here, this is accomplished by using the DP algorithm. DP line simplification algorithm is a well-known method to approximate 2D lines which was originally implemented in (Douglas and Peucker 1973). DP algorithm computes recursive construction of path hull, scaled by a tolerance factor (δ), around all points by choosing a minimum of key points. For example, in figure (8-2-A), the first path hull is constructed around five points. Each circle is drawn with radius equal to tolerance factor (δ). The approximated line in blue is constructed between two blue points which are called key points. Point 2 is outside of the path hull; therefore, the path hull should be refined. In figure (8-2-B), the second step carries on by dividing the first path hull into two path hulls which contain three key points, points 1, 3, and 5. Since all points are inside the two path hulls, the algorithm is terminated with two approximated lines.

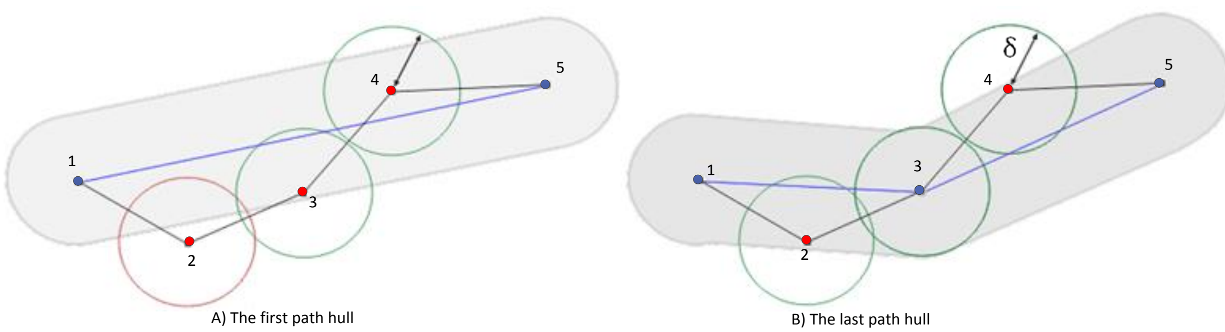


Figure 8-2: recursion steps for DP algorithm

DP algorithm is used to find the tool wear classes from collected experimental data and distinguish between adjacent cutting tool wear classes. The algorithm is used with tolerance factor (δ) equal to 0.008 in order to find the tool wear classes from tool 1 experiment with its 71 observations and tool 2 experiment with its 66 observations. All the observations of experiments 1 and 2 are shown in figure (8-4) by the points in red and in blue. The blue points are the key points. The link between those key points will be the line simplification class, approximated line. The key points are considered as indications for adjacent tool wear classes. By using Python software (Software), wear classes are obtained as: class 1 ($\mathrm{VB_B} \le 0.12$ mm), class 2 ($0.12 < \mathrm{VB_B} \le 0.1375$ mm), class 3 ($0.1375 < \mathrm{VB_B} \le 0.1749$ mm), class 4 ($0.1749 < \mathrm{VB_B} \le 0.24$ mm), class 5 ($\mathrm{VB_B} > 0.24$ mm). The mean value is taken for two blue points when the tool wear moves from class to another. For example, in class 4, starting values of tool wear are $\mathrm{VB_B} = 0.17$ mm in experiment 1 and $\mathrm{VB_B} = 0.1799$mm in experiment 2. So, the starting tool wear value for class 4 is 0.1749 mm.

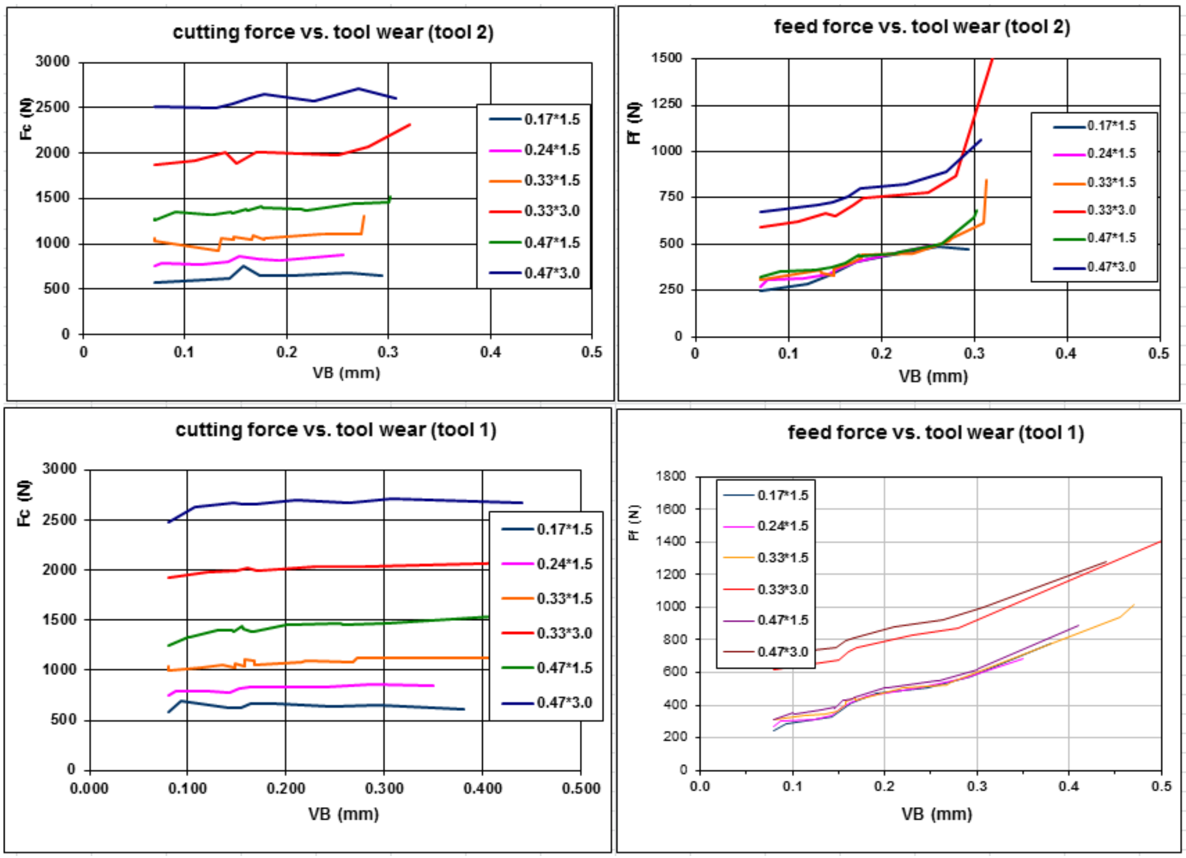


Figure 8-3: The feed forces and the cutting forces vs. the tool wear at six combinations of feed rate and depth of cut representing the six cuts.

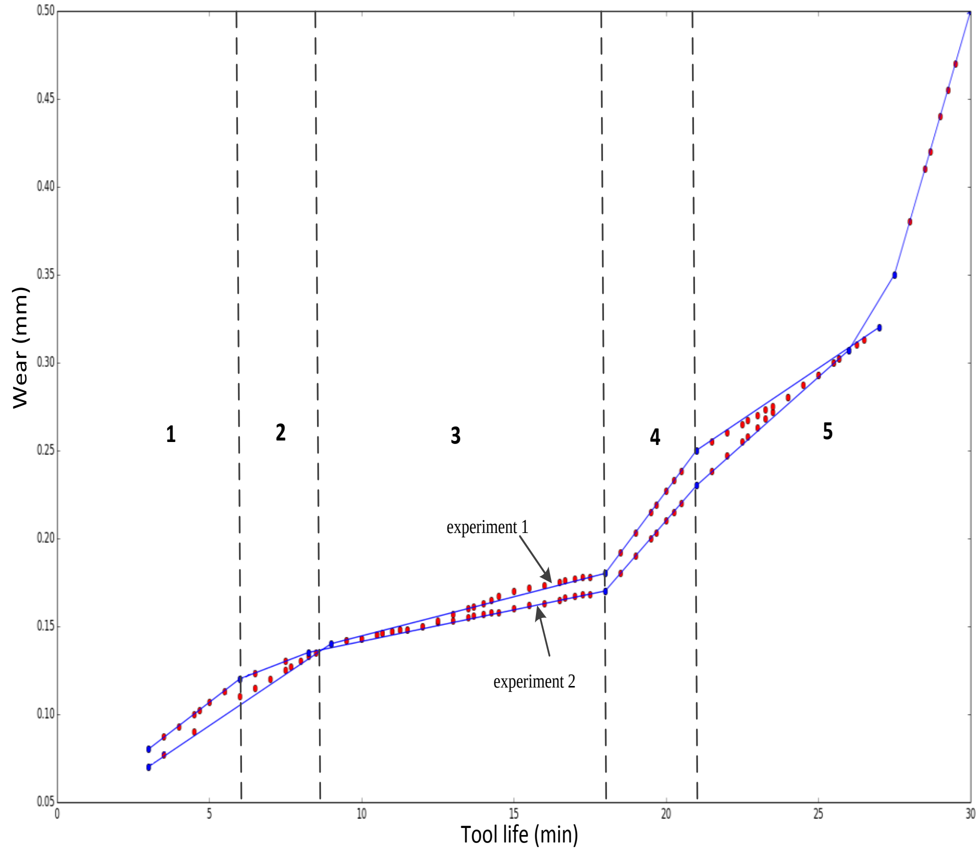


Figure 8-4: Wear classes classification

Table 8.1: Experimental data of cut 5 (tool 1, $a_p = 1.5$ mm)

Observation No	Wear (VB) mm	Class {1 to 5}	Feed (f) mm/rev	Cutting force (F_c) N	Feed force (F_f) N
1	0.080	1	0.33	1037	308
2	0.080	1	0.33	993	312
3	0.113	1	0.33	1032	334
4	0.135	2	0.33	1051	349
5	0.148	3	0.33	1032	363
6	0.148	3	0.33	1070	365
7	0.158	3	0.33	1044	405
8	0.158	3	0.33	1112	423
9	0.168	3	0.33	1089	447
10	0.168	3	0.33	1053	430
11	0.215	4	0.33	1082	499
12	0.220	4	0.33	1099	507
13	0.268	5	0.33	1088	523
14	0.272	5	0.33	1125	542
15	0.455	5	0.33	1120	941
16	0.470	5	0.33	1177	1018

8.5 Results and discussion.

Our objective is to use the experimental data in order to train LAD to identify the tool wear class by finding hidden patterns that are specific to each class. These patterns are represented in terms of the controllable and the uncontrollable machining variables, which are the feed rate (f), the feed force (F_f), and cutting force (F_C), respectively. Each pattern is a hidden rule that is found in the experimental data, and which is specific to a certain class. The data from columns 4,5, and 6, in table (8.1), for all the cuts of the two experiments, that are the 137 observations from experiments 1 and 2, are used in order to train LAD. We have five classes of cutting tool. Each observation contains variables and label. The variables are: feed rate (f), cutting force (F_c), and feed force (F_f). The label is the class number in column 3, for each observation. Table (8.2) exhibits the patterns found by the software cbmLAD (c. Software 2012).

Table 8.2: patterns found by the software cbmLAD

Class	Pattern No	Feed (f) mm/rev	Cutting force (F_c) N	Feed force (F_f) N	Class	Pattern No	Feed (f) mm/rev	Cutting force (F_c) N	Feed force (F_f) N
1	1	--	--	<313.5	3	3	--	>1984, <2025.5	--
	2	--	>929	<348		4	--	--	>374.5, <449.5
	3	--	>1733	<661		5	--	>659	>728.5, <822
	4	--	>1212	<356	4	1	--	>646	>449.5, <516.5
	5	>0.4	>2282	<728.5		2	<0.4	>2025.5,	>449.5, <848
2	1	--	>784.5	>313.5, <314.5		3	--	>2685,	>822, <881
	2	--	>1047.5, <1052	>313.5, <349.5	5	1	--	<1212	>507.5
	3	--	>1393.5	<374.5		2	--	<2685	>848
3	1	--	<830.5	>321, <449.5		3	--	<1733	>516
	2	--	<1116	>356.5, <449.5		4	--	--	>881

By analyzing the results shown in table (8.2), we were able of illustrating class identifiers of some machining variables. A variable is called class identifiers when it distinguishes the wear class by, and only, with its value, regardless other variables. For example, class 3 which is the moderate wear class is classified by the value of cutting forces only i.e. $1984 < F_c < 2025.5$. Conversely, all other patterns for other classes do not include this value. This simply means that cutting forces is a “class identifier" of class 3. Similarity, feed force can be used as class identifier of class 1 or class 5 during machining process. Feed rate cannot be used as class identifier for any class. The class identifiers are shown in table (8.3).

To measure the importance of the machining variables on tool wear classification, the frequency of their inclusion in the patterns appearing in each class, is calculated. For example, cutting force (F_c) appears in four of the five patterns of the class 1, thus the frequency is 80 %. The frequencies of all three variables; the feed rate, the feed force and the cutting force in the five classes are shown in table (8.3). According to the average of the variable frequencies, the most influential variable on tool wear classification is the feed force (F_f). And, the less influential variable on tool wear classification is the feed rate (f). This finding is conforming to the machining perspective and experimental results.

Table 8.3: machining variables’ frequencies and class identifiers in each of the five classes

Variable	Variable analysis	Class 1	Class 2	Class 3	Class 4	Class 5	Average
Feed (f)	Variable frequency	0.20	0.00	0.00	0.33	0.00	0.106
	Class identifier	--	--	--	--	--	--
Cutting force (F_c)	Variable frequency	0.80	1.00	0.80	1.00	0.75	0.87
	Class identifier	--	--	$1984 < F_c < 2025.5$	--	--	--
Feed force (F_f)	Variable frequency	1.00	1.00	0.80	1.00	1.00	0.96
	Class identifier	$F_f < 313.5$	--	$374.5 < F_f < 449.5$	--	$881 < F_f$	--

8.5.1 Validation and comparison

As many machine learning techniques, the experimental observations are divided into two sets; a training set and a testing set. The testing set is used in order to verify the quality of the learning process. The number of observations in each set is decided in a trade-off process (Russell et al. 1995). To find an accurate classifier, we need to use as much of the data as possible for the training. In contrast, to find an accurate estimate of the accuracy, we need to use as much of the data as possible for testing. For example, when we use the traditional partitioning of the dataset, we divide the dataset to equal parts, that is ½ for learning and ½ for testing; therefore, we lose 50 % of the limited number of observations. This affects negatively the learning step. In this paper, we tried different training/testing percentages. We present the results obtained when ten-fold cross validation is considered. Cross-validation technique is needed if we are to avoid peeking at the test set (Russell et al. 1995). A ten-fold cross validation procedure is a well-known technique which is performed by dividing all the observations randomly into 10 parts in which each class is represented in approximately the same proportion as in the full dataset. One part, that is 10%, is held out and is considered the testing set, T, and the learning process is trained on the remaining nine parts, that is 90%, L. The classification accuracy is calculated on the holdout, that is the testing set. The learning procedure is executed a total of 10 times with the 10 different training sets. This procedure is chosen because the largest possible volume of data is used for training. This increases the chance that the classifier found is an accurate one.

After the application of this learning and testing procedure, the average accuracy (AC) of learning is found to be equal to 85.13 %. Table (8.4) shows the results of the testing procedure applied to the training/testing dataset.

Table 8.4: the results of testing procedure

K-fold number	Training set observations	Test set observations	Accuracy
1	15 to 137	1 to 14	85.714 %
2	1 to 14 & 29 to 137	15 to 28	85.9 %
3	1 to 28 & 43 to 137	29 to 42	65.142 %
4	1 to 42 & 57 to 137	43 to 56	85.14 %
5	1 to 56 & 71 to 137	57 to 70	91.71 %
6	1 to 70 & 85 to 137	71 to 84	86.7 %
7	1 to 84 & 98 to 137	85 to 97	87.22 %
8	1 to 97 & 112 to 137	98 to 111	94.51 %
9	1 to 111 & 126 to 137	112 to 125	79.72 %
10	1 to 125	126 to 137	89.52 %
The average accuracy ($\boldsymbol{AC}$)			85.13 %

ANN is used widely as modeling technique in machining process (Shaban et al.), and tool wear detection (S Purushothaman and Srinivasa 1994; Srinivasan Purushothaman 2010; G. Wang and Cui 2013). An ANN is composed of three types of layers, namely input layer, hidden layer(s), and output layer. In figure (8-5), the input layer accepts the inputs, which are the three machining variables: feed rate (f), cutting force (F_c), and feed force (F_f). The number of neurons equal to the number of variables. The output layer has five neurons which represent the five wear classes. The number of hidden layers and its neurons depend on the nonlinearity of the model. The mathematical model of an artificial neuron's behavior mimics mathematical model of brain's activity (Russell et al. 1995). All neurons in any layer are interconnected and sending stimulating signal to the neurons of the pre and after layers through weighted links.

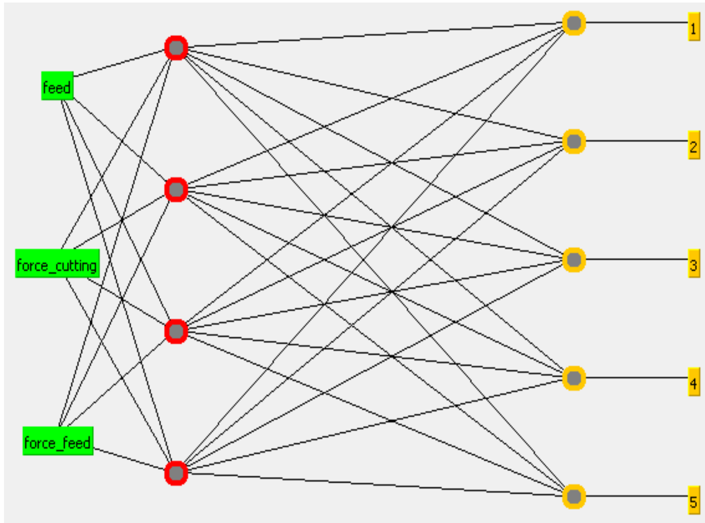


Figure 8-5: ANN architecture

After testing several neural network architecture, we retained the network architecture that gives the highest prediction accuracy in the validation test (Russell et al. 1995). The tuning parameters in neural network are the number of hidden layers, number of neurons in each layer, the learning rate, and the momentum parameter. There are no clear rules to select these parameters. Learning rate parameter and momentum parameter term take a value between [0, 1]. The value of learning rate helps the quick converge. The momentum parameter is used to update the value of a new weight by small proportion which leads to smooth searching process. The learning happens via an iterative feedback mechanism where the Back-Propagation (BP) is used to update and adjust the weights dynamically. BP is the most commonly used mechanism due to its superior strength in pattern recognition and its reasonable speed.

In this paper, we use the Weka data mining software (Hall et al. 2009) in order to find the best network structure. The common approach is to try several architecture and keep the best (Russell et al. 1995). The best network architecture which has the highest learning accuracy is found as shown in figure (8-5).It has one hidden layer of 4 neurons, a learning rate of 0.3, and a momentum of 0.2. The confusion matrix is found as follows:

$$\begin{pmatrix} G_1 & G_1^2 & G_1^3 & G_1^4 & G_1^5 \\ G_2^1 & G_2 & G_2^3 & G_2^4 & G_2^5 \\ G_3^1 & G_3^2 & G_3 & G_3^4 & G_3^5 \\ G_4^1 & G_4^2 & G_4^3 & G_4 & G_4^5 \\ G_5^1 & G_5^2 & G_5^3 & G_5^4 & G_5 \end{pmatrix} = \begin{pmatrix} 23 & 0 & 5 & 0 & 0 \\ 3 & 2 & 3 & 0 & 0 \\ 3 & 0 & 38 & 2 & 0 \\ 0 & 0 & 7 & 9 & 6 \\ 0 & 0 & 0 & 2 & 34 \end{pmatrix} ,$$

where G_m is the total number of correctly classified observations in class m, and G_m^l is the number of incorrectly classified observations which belong to class m and are classified as class l, $l \neq m$. For example, in the first row, $G_1 = 23$ is the total number of correctly classified observations in class 1, and $G_1^3 = 5$ is the number of incorrectly classified observations in class 1, and which are classified as class 3. The accuracy (AC) of learning is found to be:

$$AC = \frac{\sum_m^z G_m}{Nt} = \frac{23+2+38+9+34}{137} = 0.773$$

By using hold-out cross validation to measure the accuracy in LAD and ANN, it can be seen that the accuracy of LAD compares favourably with that of the ANN. After being trained, and by analyzing the sensors' new readings, LAD is capable of predicting the wear class based on the patterns found in each new observation that is not included in the original dataset. Based on the value $\Delta(O)$ in equation (12), the new observation belongs to the class with the highest score.

8.6 Conclusion

In this paper we have proposed a new tool wear class detective method based on pattern recognition with Logical Analysis of Data (LAD). LAD is a supervised learning data mining technique; therefore, Douglas-Peucker algorithm is used in order to find the tool wear classes from collected experimental data and distinguish between adjacent cutting tool wear classes. Based on multi-class LAD classification algorithm, the tool wear classes are defined by finding hidden patterns that are specific to each class. By analyzing the generated patterns, class identifiers of some machining variables and the influence of machining variables on tool wear classification are found. The accuracy of LAD is evaluated and validated by comparison to ANN. LAD shows better classification accuracy for tool wear.

In future work, tool wear class detective method can be used to indicate cutting tool life when cutting tool is used in rough and finish cut. The method can be improved by including additional variables, such as vibration signal, acoustic emissions, and cutting temperatures. This method can be incorporated in Computer Numerical Control (CNC) machine therefore the learning class can be done on-line.

GENERAL DISCUSSION

In this thesis, Logical Analysis of Data (LAD) is applied on different machining processes for determining the best machining conditions, detecting the tool wear, identifying the optimal replacement time for machining tools, monitoring, and controlling machining processes. LAD has demonstrated good performance and additional capabilities when it was compared to an ANN model, and to a statistical PHM. The research conclusions are as follows:

- LAD produced high accuracy in detecting the thresholds values and characteristic patterns for machining conditions even if the data are non-separable. LAD technique is used in the diagnosis of machining outcomes by comparing each incoming new measurement to the stored patterns.
- LAD is used to control the machining process by tuning autonomously the routing conditions.
- LAD is used to develop alarm system based on experimental data, under changeable machining conditions. The results show that the proposed alarm system detects the worn patterns and gives 'warning alarm' in order to replace the cutting tool at a working age that is relatively closer to the actual observed failure time than the statistical PHM.
- LAD is used to develop a new tool wear class detection method. Douglas-Peucker algorithm is used to distinguish between adjacent cutting tool wear classes. Then, LAD as a supervised learning data mining technique is used to generate decision functions.

Finally, due to the availability of sensory signals, LAD is shown to reduce the cost of machining process, by detecting accurately the time to tool replacement. The following points are my future perspectives for LAD applications in machining.

- cbmLAD and process control technique can be incorporated in computer numerical control (CNC) machine; therefore, the learning phase can be done on-line thereby eliminating the need for off-line analysis. The idea of online control of a simulated routing process can be used as a prototype for many industries' applications.

- Tool wear class detection method can be used to indicate cutting tool life when cutting tool is used in rough and finish cut. The method can be improved by including additional variables, such as vibration signal, acoustic emissions, and cutting temperatures. Moreover, tool wear classes can be identified using different methods of clustering instead of Douglas-Peucker algorithm and comparing results to DP algorithm.

- The performance of the alarm system can be improved by including additional variables, such as vibration signal, acoustic emissions, and cutting temperatures. The quality of the detected patterns will be improved, and non-pure patterns which can cover more than one class can be used, and give more details about the characteristics of LAD's patterns. Moreover, cbmLAD and tool wear monitoring alarm system will be incorporated in a computer numerical control (CNC) machine; therefore, the learning stage can be done on-line thereby eliminating the need for off-line analysis.

CONCLUSION

In this thesis, implementation of LAD on a machining process by detecting the thresholds values and characteristic patterns for machining conditions in term of uncontrollable variables using the LAD technique is presented. A simulated machining process control is implemented by using the experimental results, and LabVIEW software. The simulation model shows how LAD is used to control the routing process by tuning autonomously the routing conditions in order to always return to the machining zones defined by the positive patterns. LAD accuracy is compared to that of ANN. An on-line machining process control is developed by using the patterns that were found off-line.

A new tool wear monitoring and alarm system that is based on LAD is presented. The alarm system is a non-intrusive on-line device that measures the cutting forces and relates them to tool wear through learned patterns. It is developed during turning titanium metal matrix composites (TiMMCs).The proposed monitoring system is tested by using the experimental results obtained under sequential different machining conditions. External and internal factors that affect the turning process are taken into consideration. The system's alarm limit is validated and is compared to the limit obtained when the statistical Proportional Hazard Model (PHM) is used. The results show that the proposed system that is based on using LAD detects the worn patterns and gives a more accurate alarm for cutting tool replacement. We also show how to exploit condition monitoring data in machining operation in order to extract intelligent knowledge, and use this knowledge to determine the tool replacement time. Finally, a new tool wear multi-class detection method is presented. Based on experimental data, tool wear classes are defined using Douglas-Peucker algorithm. LAD is then used as a machine learning, pattern recognition technique for double objective of detecting the present tool wear class based on the recent sensors' readings of the time dependant machining variables, and deriving novel information about the role of machining variables by doing patterns analysis. The accuracy of LAD is compared to ANN.

It is expected that an analysis tool such as LAD will help in blazing a new trail in machining processes by using state of the art techniques in order to significantly reduce the cost of machining process. I hope the results of my work will have an impact in the future of optimization of machining processes.

REFERENCES

Achiche, S., Balazinski, M., Baron, L., & Jemielniak, K. (2002). Tool wear monitoring using genetically-generated fuzzy knowledge bases. *Engineering Applications of Artificial Intelligence, 15*(3-4), 303-314, doi:10.1016/s0952-1976(02)00071-4.

Alexe, G., Alexe, S., Bonates, T., & Kogan, A. (2007). Logical analysis of data – the vision of Peter L. Hammer. *Annals of Mathematics and Artificial Intelligence, 49*(1), 265-312, doi:10.1007/s10472-007-9065-2.

Attia, M. H., Sadek, A., and Meshreki, M (2011.). High Speed Machining of Composite Materials ", Chapter in the book on Technology for Composite Materials: Principles and Practice. *(ed. H. Hocheng), Woodhead Publishing Limited, Cambridge, UK*.

Aven, T., & Bergman, B. (1986). Optimal replacement times: a general set-up. *Journal of Applied Probability*, 432-442.

Azouzi, R., & Guillot, M. (1997). On-line prediction of surface finish and dimensional deviation in turning using neural network based sensor fusion. *International Journal of Machine Tools and Manufacture*, *37*(9), 1201-1217.

Balazinski, M., Czogala, E., Jemielniak, K., & Leski, J. (2002). Tool condition monitoring using artificial intelligence methods. *Engineering Applications of Artificial Intelligence, 15*(1), 73-80, doi:10.1016/s0952-1976(02)00004-0.

Balazinski, M., & Mpako, C. (2000). Improvement of tool life through the use of two discrete feed rates during machining of 4140 steel. *Machining Science and Technology, 4*(1), 1-13.

Banjevic, D., Jardine, A., Makis, V., & Ennis, M. (2001). A control-limit policy and software for condition-based maintenance optimization. *INFOR-OTTAWA-, 39*(1), 32-50.

Benardos, P., & Vosniakos, G. (2002). Prediction of surface roughness in CNC face milling using neural networks and Taguchi's design of experiments. *Robotics and Computer-Integrated Manufacturing, 18*(5), 343-354.

Bennane, A., & Yacout, S. (2012). LAD-CBM; new data processing tool for diagnosis and prognosis in condition-based maintenance. *Journal of Intelligent Manufacturing, 23*(2), 265-275.

Bergman, B. (1978). Optimal replacement under a general failure model. *Advances in Applied Probability*, 431-451.

Bores, E., Hammer, P. L., Ibaraki, T., Kogan, A., Mayoraz, E., & Muchnik, I. (2000). An implementation of logical analysis of data. *Knowledge and Data Engineering, IEEE Transactions on, 12*(2), 292-306.

Bustillo, A., & Correa, M. (2012). Using artificial intelligence to predict surface roughness in deep drilling of steel components. *Journal of Intelligent Manufacturing, 23*(5), 1893-1902.

Byrne, G., Dornfeld, D., Inasaki, I., Ketteler, G., König, W., & Teti, R. (1995). Tool condition monitoring (TCM)—the status of research and industrial application. *CIRP Annals-Manufacturing Technology, 44*(2), 541-567.

Çaydaş, U., & Ekici, S. (2012). Support vector machines models for surface roughness prediction in CNC turning of AISI 304 austenitic stainless steel. *Journal of Intelligent Manufacturing, 23*(3), 639-650.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2011). SMOTE: synthetic minority over-sampling technique. *arXiv preprint arXiv:1106.1813*.

Chen, W.-C. (1997). Some experimental investigations in the drilling of carbon fiber-reinforced plastic (CFRP) composite laminates. *International Journal of Machine Tools and Manufacture, 37*(8), 1097-1108.

Choudhary, A. K., Harding, J. A., & Tiwari, M. K. (2009). Data mining in manufacturing: a review based on the kind of knowledge. *Journal of Intelligent Manufacturing, 20*(5), 501-521.

Coker, S. A., & Shin, Y. C. (1996). In-process control of surface roughness due to tool wear using a new ultrasonic system. *International Journal of Machine Tools and Manufacture, 36*(3), 411-422.

Crama, Y., Hammer, P. L., & Ibaraki, T. (1988). Cause-effect relationships and partially defined Boolean functions. *Annals of Operations Research, 16*(1), 299-325.

Damodarasamy, S., & Raman, S. (1993). An inexpensive system for classifying tool wear states using pattern recognition. *Wear, 170*(2), 149-160.

Davim, J. P., & Reis, P. (2005). Damage and dimensional precision on milling carbon fiber-reinforced plastics using design experiments. *Journal of Materials Processing Technology, 160*(2), 160-167.

Ding, F., & He, Z. (2011). Cutting tool wear monitoring for reliability analysis using proportional hazards model. *The International Journal of Advanced Manufacturing Technology, 57*(5-8), 565-574.

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization, 10*(2), 112-122.

Du, S., Lv, J., & Xi, L. (2012). A robust approach for root causes identification in machining processes using hybrid learning algorithm and engineering knowledge. *Journal of Intelligent Manufacturing, 23*(5), 1833-1847.

Elliott, C., Vijayakumar, V., Zink, W., & Hansen, R. (2007). National instruments LabVIEW: a programming environment for laboratory automation and measurement. *Journal of the Association for Laboratory Automation, 12*(1), 17-24.

Ertunc, H. M., & Oysu, C. (2004). Drill wear monitoring using cutting force signals. *Mechatronics, 14*(5), 533-548.

. EXAKT help Version 4.20.1

(2007). Optimal Maintenance Decisions (OMDEC) Inc.

Ferreira, J., Coppini, N., & Miranda, G. (1999). Machining optimisation in carbon fibre reinforced composite materials. *Journal of Materials Processing Technology, 92*, 135-140.

Gray, A. E., Seidmann, A., & Stecke, K. E. (1993). A synthesis of decision models for tool management in automated manufacturing. *Management science, 39*(5), 549-567.

Haber, R. E., Haber, R., Alique, A., & Ros, S. (2002). Application of knowledge-based systems for supervision and control of machining processes. *Handbook of software engineering and knowledge engineering, 2*, 327-362.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10-18.

Hammer, P. L. Partially defined Boolean functions and cause-effect relationships. In *International conference on multi-attribute decision making via or-based expert systems, 1986*

Hammer, P. L., & Bonates, T. O. (2006). Logical analysis of data—an overview: from combinatorial optimization to medical applications. *Annals of Operations Research, 148*(1), 203-225.

Han, J., Kim, N., Yum, B.-J., & Jeong, M. K. (2011). Pattern selection approaches for the logical analysis of data considering the outliers and the coverage of a pattern. *Expert Systems with Applications, 38*(11), 13857-13862.

Hansen, P., & Meyer, C. (2011). A new column generation algorithm for Logical Analysis of Data. *Annals of Operations Research, 188*(1), 215-249.

Herrera, J. F. A., & Subasi, M. M. R utcor Research R eport.

Ho-Cheng, H., & Dharan, C. (1990). Delamination during drilling in composite laminates. *Journal of Engineering for Industry(Transactions of the ASME), 112*(3), 236-239.

Huang, P. B. (2014). An intelligent neural-fuzzy model for an in-process surface roughness monitoring system in end milling operations. *Journal of Intelligent Manufacturing*, 1-12.

Huang, Y., & Liang, S. Y. (2005). Modeling of cutting forces under hard turning conditions considering tool wear effect. *Journal of manufacturing science and engineering, 127*(2), 262-270.

Hui, Y., & Leung, L. (1994). Optimal economic tool regrinding with Taguchi's quality loss function. *The Engineering Economist, 39*(4), 313-331.

Jeang, A. (1998). Reliable tool replacement policy for quality and cost. *European journal of operational research, 108*(2), 334-344.

Jemielniak, K. (1999). Commercial tool condition monitoring systems. *The International Journal of Advanced Manufacturing Technology, 15*(10), 711-721.

Kalbfleisch, J. D., & Prentice, R. L. (2011). *The statistical analysis of failure time data* (Vol. 360): John Wiley & Sons.

Kang, J., Kang, N., Feng, C.-j., & Hu, H.-y. Research on tool failure prediction and wear monitoring based hmm pattern recognition theory. In *Wavelet Analysis and Pattern*

Recognition, 2007. ICWAPR'07. International Conference on, 2007 (Vol. 3, pp. 1167-1172): IEEE

Kannan, S., Balazinski, M., & Kishawy, H. (2006). Flank wear progression during machining metal matrix composites. *Journal of manufacturing science and engineering, 128*(3), 787-791.

Klim, Z., Ennajimi, E., Balazinski, M., & Fortin, C. (1996). Cutting tool reliability analysis for variable feed milling of 17-4PH stainless steel. *Wear, 195*(1), 206-213.

Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence, 1995* (Vol. 14, pp. 1137-1145): Lawrence Erlbaum Associates Ltd

Landers, R. G., Ulsoy, A. G., & Furness, R. J. (2002). Process monitoring and control of machining operations. *The Mechanical Systems Design Handbook*.

Li, B. (2012). A review of tool wear estimation using theoretical analysis and numerical simulation technologies. *International Journal of Refractory Metals and Hard Materials, 35*, 143-151.

Liang, S. Y., Hecker, R. L., & Landers, R. G. (2004). Machining process monitoring and control: the state-of-the-art. *Journal of manufacturing science and engineering, 126*(2), 297-310.

Lin, T., & Shyu, R.-F. (2000). Improvement of tool life and exit burr using variable feeds when drilling stainless steel with coated drills. *The International Journal of Advanced Manufacturing Technology, 16*(5), 308-313.

Linderoth, J. T., & Lodi, A. (2011). MILP software. *Wiley Encyclopedia of Operations Research and Management Science*.

Liu, H. (1997). *Modeling and optimal control of deteriorating production processes*. University of Toronto,

Liu, H., & Makis, V. (1996). Cutting-tool reliability assessment in variable machining conditions. *Reliability, IEEE Transactions on, 45*(4), 573-581.

Liu, P. H., Makis, V., & Jardine, A. K. (2001). Scheduling of the optimal tool replacement times in a flexible manufacturing system. *IIE Transactions, 33*(6), 487-495.

Makis, V. (1995). Optimal replacement of a tool subject to random failure. *International journal of production economics, 41*(1), 249-256.

Makis, V., & Jardine, A. K. (1992). Optimal replacement in the proportional hazards model. *Infor, 30*(1), 172-183.

Mayoraz, E., & Moreira, M. On the decomposition of polychotomies into dichotomies. In*, 1997* (pp. 219-226): MORGAN KAUFMANN PUBLISHERS, INC.

Mayoraz, E., & Moreira, M. (1999). Combinatorial approach for data binarization. In *Principles of data mining and knowledge discovery* (pp. 442-447): Springer.

Mazzuchi, T. A., & Soyer, R. (1989). Assessment of machine tool reliability using a proportional hazards model. *Naval Research Logistics (NRL), 36*(6), 765-777.

Meshreki, M., Sadek, A., & Attia, M. H. High Speed Routing of Woven Carbon Fiber Reinforced Epoxy Laminates. In *Proceedings of the ASME 2012 International Mechanical Engineering Congress & Exposition, Houston, Texas, USA, 2012*

Montgomery, D. C. (2007). *Introduction to statistical quality control*: John Wiley & Sons.

Moreira, L. M. (2000). *The use of Boolean concepts in general classification contexts*. Universidade do Minho, Portugal,

Mortada, Yacout, S., & Lakis, A. (2011). *Applicability and interpretability of logical analysis of data in condition based maintenance*. NR79949, Ecole Polytechnique, Montreal (Canada), Canada.

Mortada, M.-A., Yacout, S., & Lakis, A. (2011). Diagnosis of rotor bearings using logical analysis of data. *Journal of Quality in Maintenance Engineering, 17*(4), 371-397, doi:10.1108/13552511111180186.

Mortada, M.-A., Yacout, S., & Lakis, A. (2013). Fault diagnosis in power transformers using multi-class logical analysis of data. *Journal of Intelligent Manufacturing*, 1-11.

Mortada, M. A., Carroll Iii, T., Yacout, S., & Lakis, A. (2009). Rogue components: their effect and control using logical analysis of data. *Journal of Intelligent Manufacturing*, 1-14.

Purushothaman, S. (2010). Tool wear monitoring using artificial neural network based on extended Kalman filter weight updation with transformed input patterns. *Journal of Intelligent Manufacturing, 21*(6), 717-730.

Purushothaman, S., & Srinivasa, Y. (1994). A back-propagation algorithm applied to tool wear monitoring. *International Journal of Machine Tools and Manufacture, 34*(5), 625-631.

Rahman, M., Ramakrishna, S., Prakash, J., & Tan, D. (1999). Machinability study of carbon fiber reinforced composite. *Journal of Materials Processing Technology, 89*, 292-297.

Rangwala, S. S. (1988). *Machining process characterization and intelligent tool condition monitoring using acoustic emission signal analysis*. University of California, Berkeley,(Dec,

Rawat, S., & Attia, H. (2009a). Characterization of the dry high speed drilling process of woven composites using Machinability Maps approach. *CIRP Annals-Manufacturing Technology, 58*(1), 105-108.

Rawat, S., & Attia, H. (2009b). Wear mechanisms and tool life management of WC–Co drills during dry high speed drilling of woven carbon fibre composites. *Wear, 267*(5), 1022-1030.

Ren, Q., Balazinski, M., Baron, L., & Jemielniak, K. Tool condition monitoring using the TSK fuzzy approach based on subtractive clustering method. In *21st International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2008, June 18, 2008 - June 20, 2008, Wroclaw, Poland, 2008* (Vol. 5027 LNAI, pp. 52-61, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)): Springer Verlag. doi:10.1007/978-3-540-69052-8_6.

Russell, S. J., Norvig, P., Canny, J. F., Malik, J. M., & Edwards, D. D. (1995). *Artificial intelligence: a modern approach* (Vol. 74): Prentice hall Englewood Cliffs.

Ryoo, H. S., & Jang, I. Y. (2009). Milp approach to pattern generation in logical analysis of data. *Discrete Applied Mathematics, 157*(4), 749-761.

Sakharov, G., Ilinykh, V., & Konyukhov, V. Y. (1990). Improvement of fastening elements in an assembled cutting tool. *Soviet Engineering Research, 10*, 102.

Salamanca, D., & Yacout, S. (2007). Condition based maintenance with logical analysis of data. *7e Congrès International de génie industriel*.

Shaban, Y., Aramesh, M., Yacout, S., Balazinski, M., Attia, H., & Kishawy, H. Optimal replacement of tool during turning titanium metal matrix composites. In *Proceedings of the 2014 Industrial and Systems Engineering Research Conference, Montreal,*

Shaban, Y., Meshreki, M., Yacout, S., Balazinski, M., & Attia, H. Process control based on pattern recognition for routing carbon fiber reinforced polymer. *Journal of Intelligent Manufacturing*, 1-15.

Sharma, V. S., Dhiman, S., Sehgal, R., & Sharma, S. (2008). Estimation of cutting forces and surface roughness for hard turning using neural networks. *Journal of Intelligent Manufacturing, 19*(4), 473-483.

Shi, D., & Gindy, N. N. (2007). Tool wear predictive model based on least squares support vector machines. *Mechanical Systems and Signal Processing, 21*(4), 1799-1814.

Sick, B. (2002). On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. *Mechanical Systems and Signal Processing, 16*(4), 487-546.

Software Python Language Reference, version 2.7. Available at http://www.python.org.

Software, c. (2012). Patent Cooperation Treaty PCT/CA2011/000876, No. Wo 2012/00984.

Tail, M., Yacout, S., & Balazinski, M. (2010). Replacement time of a cutting tool subject to variable speed. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 224*(3), 373-383, doi:10.1243/09544054jem1693.

Tansel, I., Demetgul, M., Bickraj, K., Kaya, B., & Ozcelik, B. (2013). Basic computational tools and mechanical hardware for torque-based diagnostic of machining operations. *Journal of Intelligent Manufacturing, 24*(1), 147-161.

Taylor, F. W. (1907). On the art of cutting metals.

Teti, R. (2002). Machining of composite materials. *CIRP Annals-Manufacturing Technology, 51*(2), 611-634.

Tobon-Mejia, D., Medjaher, K., & Zerhouni, N. (2012). CNC machine tool's wear diagnostic and prognostic by using dynamic Bayesian networks. *Mechanical Systems and Signal Processing, 28*, 167-182.

Tomac, N., Tannessen, K., & Rasch, F. O. (1992). Machinability of particulate aluminium matrix composites. *CIRP Annals-Manufacturing Technology, 41*(1), 55-58.

Vagnorius, Z., Rausand, M., & Sørby, K. (2010). Determining optimal replacement time for metal cutting tools. *European journal of operational research, 206*(2), 407-416.

Wang, G., & Cui, Y. (2013). On line tool wear monitoring based on auto associative neural network. *Journal of Intelligent Manufacturing, 24*(6), 1085-1094.

Wang, H., & Huang, Q. (2006). Error cancellation modeling and its application to machining process control. *IIE Transactions, 38*(4), 355-364.

Wang, W. (2004). Proportional hazards regression models with unknown link function and time-dependent covariates. *Statistica Sinica, 14*(3), 885-906.

Wang, W., & Hu, C. (2006). Proportional Hazards Regression Models. In *Springer Handbook of Engineering Statistics* (pp. 387-396): Springer.

Witten, I. H., & Frank, E. (2011). *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann.

Wolpert, D. H. (1996). The lack of a priori distinctions between learning algorithms. *Neural computation, 8*(7), 1341-1390.

Yacout, S. Fault detection and diagnosis for condition based maintenance using the logical analysis of data. In *40th International Conference on Computers and Industrial Engineering, , Japan, 2010*: IEEE Computer Society. doi:10.1109/iccie.2010.5668357.

Yacout, S., Meshreki, M., & Attia, H. Monitoring and control of machining process by data mining and pattern recognition. In *2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS),, 2012* (pp. 106-113): IEEE Computer Society

Zhang, J. Z., Chen, J. C., & Kirby, E. D. (2007). The development of an in-process surface roughness adaptive control system in turning operations. *Journal of Intelligent Manufacturing, 18*(3), 301-311.

Zuperl, U., Cus, F., & Reibenschuh, M. (2012). Modeling and adaptive force control of milling by using artificial techniques. *Journal of Intelligent Manufacturing, 23*(5), 1805-1815.